

NASA Contractor Report 187443

ICASE INTERIM REPORT 14

SPECTRAL METHODS FOR TIME DEPENDENT PROBLEMS

Eitan Tadmor

(NASA-CR-187443) SPECTRAL METHODS FOR TIME
DEPENDENT PROBLEMS Interim Report (ICASE)
67 p CSCL 12A

N91-10064

Unclass

63/64 0310075

NASA Contract No. NAS1-18605
September 1990

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the Universities Space Research Association



National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665-5225



ICASE INTERIM REPORTS

ICASE has introduced a new report series to be called ICASE Interim Reports. The series will complement the more familiar blue ICASE reports that have been distributed for many years. The blue reports are intended as preprints of research that has been submitted for publication in either refereed journals or conference proceedings. In general, the green Interim Report will not be submitted for publication, at least not in its printed form. It will be used for research that has reached a certain level of maturity but needs additional refinement, for technical reviews or position statements, for bibliographies, and for computer software. The Interim Reports will receive the same distribution as the ICASE Reports. They will be available upon request in the future, and they may be referenced in other publications.

Robert G. Voigt
Director

SPECTRAL METHODS FOR TIME DEPENDENT PROBLEMS

Eitan Tadmor¹

School of Mathematical Sciences, Tel-Aviv University

ABSTRACT

This short review on spectral approximations for time-dependent problems consists of three parts. In part I we discuss some basic ingredients from the spectral Fourier and Chebyshev approximation theory. Part II contains a brief survey on hyperbolic and parabolic time-dependent problems which are dealt with both the energy method and the related Fourier analysis. In part III we combine the ideas presented in the first two parts, in our study of accuracy stability and convergence of the spectral Fourier approximation to time-dependent problems.

Lecture notes for the

Nordic Summerschool on Numerical Methods in Fluid Dynamics

August 20-25, 1990
Sydkoster, SWEDEN

¹Research was supported in part by NASA Contract No. NAS1-18107 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665.

TABLE OF CONTENTS

1. SPECTRAL APPROXIMATION THEORY	1
1.1. The Periodic Problem - Fourier Approximation	1
1.2. The Pseudospectral (Collocation) Approximation	8
1.3. Spectral and Pseudospectral Approximations - Exponential Accuracy	17
1.4. The Non-Periodic Problem - Chebyshev Approximation	19
2. TIME DEPENDENT PROBLEMS	29
2.1. Initial Value Problems of Hyperbolic Type	29
2.2. Initial Value Problems of Parabolic Type	36
2.3. Well-posed Time-Dependent Problems	38
3. THE FOURIER METHOD FOR HYPERBOLIC AND PARABOLIC EQUATIONS	40
3.1. The Spectral Approximation	40
3.2. The Pseudospectral Approximation	50
3.3. Skew-Symmetric Differencing	54
3.4. Smoothing	55
Appendix. Fourier Collocation with Even Number of Gridpoints	62

1. SPECTRAL APPROXIMATION THEORY

1.1. The Periodic Problem - Fourier Approximation

Consider the first order Sturm-Liouville (SL) problem

$$(1.1.1) \quad \frac{d}{dx}\phi = \lambda\phi(x), \quad 0 \leq x \leq 2\pi,$$

augmented by periodic boundary conditions

$$(1.1.2) \quad \phi(0) = \phi(2\pi).$$

It has an infinite sequence of eigenvalues, $\lambda_k = ik$, with the corresponding eigenfunctions $\phi_k(x) = e^{ikx}$. Thus, $(\lambda_k = ik, \phi_k = e^{ikx})$ are the eigenpairs of the differentiation operator $D \equiv \frac{d}{dx}$ in $L^2[0, 2\pi)$, and they form a complete system in this space. This system is complete in the following sense. Let $L^2[0, 2\pi)$ be induced with the usual Euclidean inner product

$$(1.1.3) \quad (w_1(x), w_2(x)) \equiv \int_0^{2\pi} w_1(x) \overline{w_2(x)} dx.$$

Note that $\phi_k(x)$ are orthogonal with respect to this inner product, for

$$(1.1.4) \quad (e^{ikx}, e^{ijx}) = \begin{cases} 0 & j \neq k, \\ \|e^{ikx}\|^2 = 2\pi & j = k. \end{cases}$$

Let $w(x) \in L^2[0, 2\pi)$ be associated with its spectral representation in this system, i.e., the Fourier expansion

$$(1.1.5) \quad w(x) \sim \sum_{k=-\infty}^{\infty} \hat{w}(k) \phi_k(x), \quad \hat{w}(k) = \frac{(w, \phi_k)}{\|\phi_k\|^2},$$

or equivalently,

$$(1.1.6) \quad w(x) \sim \sum_{k=-\infty}^{\infty} \hat{w}(k) e^{ikx}, \quad \hat{w}(k) = \frac{1}{2\pi} \int_{\xi=0}^{2\pi} w(\xi) e^{-ik\xi} d\xi.$$

The truncated Fourier expansion

$$(1.1.7) \quad S_N w \equiv \sum_{k=-N}^N \hat{w}(k) e^{ikx},$$

denotes the spectral-Fourier projection of $w(x)$ into π_N -the space of trigonometric polynomials of degree $\leq N$:

$$\begin{aligned}
 S_N w &= \hat{w}(0) + \sum_{k=1}^N [\hat{w}(k)e^{ikx} + \hat{w}(-k)e^{-ikx}] \\
 (1.1.8) \quad &= \hat{w}(0) + \sum_{k=1}^N [\hat{w}(k) + \hat{w}(-k)] \cos kx + i[\hat{w}(k) - \hat{w}(-k)] \sin kx \\
 &= \sum_{k=0}^N \hat{a}_k \cos kx + \hat{b}_k \sin kx;
 \end{aligned}$$

here \hat{a}_k and \hat{b}_k are the usual Fourier coefficients given by

$$\begin{aligned}
 (1.1.9) \quad \hat{a}_k &= \hat{w}(k) + \hat{w}(-k) = \frac{1}{\pi} \int_0^{2\pi} w(\xi) \cos k\xi d\xi, \\
 \hat{b}_k &= i[\hat{w}(k) - \hat{w}(-k)] = \frac{1}{\pi} \int_0^{2\pi} w(\xi) \sin k\xi d\xi.
 \end{aligned}$$

Since $w - S_N w$ is orthogonal to the π_N -space:

$$(1.1.10) \quad (w - S_N w, e^{ikx}) = 2\pi \hat{w}(k) - 2\pi \hat{w}(k) = 0, \quad |k| \leq N,$$

it follows that for any $p_N \in \pi_N$ we have (see Figure 1)

$$(1.1.11) \quad \|w - p_N\|^2 = \|w - S_N w\|^2 + \|S_N w - p_N\|^2.$$

Hence, $S_N w$ solves the least-squares problem

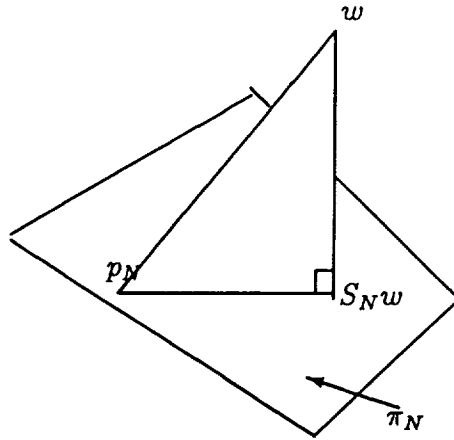


Figure 1:

$$(1.1.12) \quad \|w - S_N w\| = \text{Min}_{p_N \in \pi_N} \|w - p_N\|$$

i.e., $S_N w$ is the best least-squares approximation to w . Moreover, (1.1.11) with $p_N = 0$ yields

$$(1.1.13) \quad \|S_N w\|^2 = \|w\|^2 - \|w - S_N w\|^2 \leq \|w\|^2$$

and by letting $N \rightarrow \infty$ we arrive at Bessel's inequality

$$(1.1.14) \quad 2\pi \sum_{k=-\infty}^{\infty} |\hat{w}(k)|^2 \equiv \sum_{k=-\infty}^{\infty} |\hat{w}(k)|^2 \|\phi_k\|^2 \leq \|w\|^2.$$

Remark: An immediate consequence of (1.1.14) is the Riemann-Lebesgue lemma, asserting that

$$\hat{w}(k) = \frac{1}{2\pi} \int_0^{2\pi} w(\xi) e^{-ik\xi} d\xi \xrightarrow[k \rightarrow \infty]{} 0, \quad \text{for any } w \in L^2[0, 2\pi).$$

The system $\{\phi_k = e^{ikx}\}$ is complete in the sense that for any $w(x) \in L^2[0, 2\pi)$ we have Parseval equality:

$$(1.1.15) \quad 2\pi \sum_{k=-\infty}^{\infty} |\hat{w}(k)|^2 \equiv \sum_{k=-\infty}^{\infty} |\hat{w}(k)|^2 \|\phi_k\|^2 = \|w\|^2,$$

which in view of (1.1.13), is the same as

$$(1.1.16) \quad \lim_{N \rightarrow \infty} \|S_N w \equiv \sum_{-N}^N \hat{w}(k) e^{ikx} - w(x)\| = 0.$$

The last equality establishes the L^2 convergence of the spectral-Fourier projection, $S_N w(x)$, to $w(x)$, whose difference can be upper bounded by the following

Error Estimate:

$$\|w - S_N w\|^2 = \|w\|^2 - \|S_N w\|^2 = \sum_{|k| > N} |\hat{w}(k)|^2 \|\phi_k\|^2 = 2\pi \sum_{|k| > N} |\hat{w}(k)|^2.$$

We observe that the RHS tends to zero as a tail of a converging sequence, i.e.,

$$(1.1.17) \quad \int_0^{2\pi} |w(x) - \sum_{k=-N}^N \hat{w}(k) e^{ikx}|^2 dx = 2\pi \sum_{|k| > N} |\hat{w}(k)|^2 \xrightarrow[N \rightarrow \infty]{} 0.$$

The last equality tells us that the convergence rate depends on how fast the Fourier coefficients, $\hat{w}(k)$, decay to zero, and we shall quantify this in a more precise way below.

Remark (1.1.17) yields uniform a.e., convergence for subsequences; in fact one can show

$$(1.1.18) \quad \text{a.e. } \lim_{p \rightarrow \infty} |w(x) - S_{N_p} w(x)| = 0, \quad \inf_p \frac{N_{p+1}}{N_p} > 1.$$

In fact, $w(x) = \text{a.e.} \lim_{N \rightarrow \infty} S_N w(x)$ for all $w \in L^2[0, 2\pi]$, but a.e. convergence may fail if $w(\cdot)$ is only $L^1[0, 2\pi]$ -integrable.

Yet, if we agree to assume sufficient smoothness, we find the convergence of spectral-Fourier projection to be very rapid, both in the L^2 and the pointwise sense. To this we proceed as follows.

Define the Sobolev space $H^s[0, 2\pi)$ consisting of 2π -periodic functions which their first s -derivatives are L^2 -integrable; set the inner product

$$(1.1.19) \quad (w_1, w_2)_{H^s} = \sum_{p=0}^s \int_0^{2\pi} D^p w_1(x) \overline{D^p w_2(x)} dx.$$

The essential ingredient here is that the system $\{e^{ikx}\}$ – which was already shown to be complete in $L^2[0, 2\pi) \equiv H^0[0, 2\pi)$, is also a complete system in $H^s[0, 2\pi)$ for any $s \geq 0$. For orthogonality we have

$$(1.1.20) \quad (e^{ikx}, e^{ijx})_{H^s} = \begin{cases} 0 & j \neq k, \\ 2\pi \sum_{p=0}^s k^{2p} & j = k. \end{cases}$$

The Fourier expansion now reads

$$(1.1.21) \quad w(x) \sim \sum_{k=-\infty}^{\infty} \hat{w}_s(k) e^{ikx}$$

where the Fourier coefficients, $\hat{w}_s(k)$, are given by

$$(1.1.22) \quad \hat{w}_s(k) = \frac{(w(x), e^{ikx})_{H^s}}{(e^{ikx}, e^{ikx})_{H^s}}.$$

We integrate by parts and use periodicity to obtain

$$\begin{aligned} (w(x), e^{ikx})_{H^s} &= \sum_{p=0}^s \int_0^{2\pi} D^p w(x) \overline{D^p e^{ikx}} dx = \\ &= \sum_{p=0}^s (-1)^p \int_0^{2\pi} w(x) \overline{D^{2p} e^{ikx}} dx \\ &= \sum_{p=0}^s (-1)^p (-ik)^{2p} \int_0^{2\pi} w(\xi) e^{-ik\xi} d\xi \end{aligned}$$

and together with (1.1.20) we recover the usual Fourier expansion we had before, namely

$$(1.1.23) \quad \hat{w}_s(k) \equiv \hat{w}(k) = \frac{1}{2\pi} \int_{\xi=0}^{2\pi} w(\xi) e^{-ik\xi} d\xi.$$

The completion of $\{e^{ikx}\}$ in $H^s[0, 2\pi)$ gives us the Parseval equality, compare (1.1.15)

$$(1.1.24) \quad \|w - S_N w\|_{H^s}^2 = \sum_{|k| > N} |\hat{w}_s(k)|^2 \|e^{ikx}\|_{H^s}^2 = \sum_{|k| > N} [|\hat{w}(k)|^2 \cdot 2\pi \sum_{p=0}^s k^{2p}] \geq$$

$$\sum_{p=0}^s N^{2p} \cdot 2\pi \cdot \sum_{|k| > N} |\hat{w}(k)|^2 = \sum_{p=0}^s N^{2p} \cdot \|w - S_N w\|^2.$$

Since

$$(1.1.25) \quad \text{Const}_1(1 + N^2)^{s/2} \leq \left(\sum_{p=0}^s N^{2p} \right)^{\frac{1}{2}} \leq \text{Const}_2(1 + N^2)^{\frac{s}{2}},$$

we conclude from (1.1.24), that for any $w \in H^s[0, 2\pi)$ we have

$$(1.1.26) \quad \|w - S_N w\| \leq \text{Const}_s \cdot \frac{1}{N^s}, \quad w \in H^s[0, 2\pi).$$

Note that $\text{Const}_s = \text{Const}_1 \cdot \|w - S_N w\|_{H^s} \xrightarrow{N \rightarrow \infty} 0$. This kind of estimate is usually referred to by saying that the Fourier expansion has spectral accuracy, i.e., the error tends to zero faster than any fixed power of N , and is restricted only by the global smoothness of $w(x)$.

We note that as before, this kind of behavior is linked directly to the spectral decay of the Fourier coefficients in this case. Indeed, by Cauchy-Schwartz inequality

$$(1.1.27) \quad |\hat{w}(k)| = |\hat{w}_s(k)| \leq \frac{\|w\|_{H^s} \cdot \|e^{ikx}\|_{H^s}}{\|e^{ikx}\|_{H^s}^2} \leq \frac{1}{(2\pi \sum_{p=0}^s k^{2p})^{\frac{1}{2}}} \|w\|_{H^s}$$

$$\leq \text{Const} \cdot \frac{1}{(1 + |k|^2)^{\frac{s}{2}}}.$$

In fact more is true. By Parseval equality

$$\|w\|_{H^s}^2 = \sum_{k=-\infty}^{\infty} |\hat{w}(k)|^2 \|e^{ikx}\|_{H^s}^2 = 2\pi \sum_{k=-\infty}^{\infty} \left(\sum_{p=0}^s k^{2p} \right) |\hat{w}(k)|^2,$$

and hence by the Reimann-Lebesgue lemma, the product $(1 + |k|^2)^{\frac{s}{2}} |\hat{w}(k)|$ is not only bounded (as asserted in (1.1.27), but in fact it tends to zero,

$$(1 + |k|^2)^{\frac{s}{2}} |\hat{w}(k)| \xrightarrow{k \rightarrow \infty} 0.$$

Thus, $\hat{w}(k)$ tends to zero faster than $|k|^{-s}$ for all $w(x) \in H^s$. This yields spectral convergence, for

$$\|w - S_N w\|^2 = 2\pi \sum_{|k| > N} |\hat{w}(k)|^2 \leq \text{Const} \cdot \sum_{|k| > N} \frac{1}{(1 + |k|^2)^s} \leq \text{Const} \cdot \frac{1}{N^{2s-1}}$$

i.e., we get slightly less than (1.1.26),

$$\|w - S_N w\| \leq \text{Const.} \frac{1}{N^{s-\frac{1}{2}}} \xrightarrow{N \rightarrow \infty} 0 \quad s \geq 1.$$

Moreover, there is a rapid convergence for derivatives as well. Indeed, if $w(x) \in H^s[0, 2\pi]$ then for $0 < \sigma < s$ we have

$$\begin{aligned} \|w - S_N w\|_{H^\sigma}^2 &= \sum_{|k| > N} (2\pi \sum_{p=0}^{\sigma} k^{2p}) |\hat{w}(k)|^2 \\ &\leq \text{Const.} \sum_{|k| > N} (1 + |k|^2)^\sigma |\hat{w}(k)|^2 \\ &\leq \text{Const.} \sum_{|k| > N} \frac{(1 + |k|^2)^s}{(1 + N^2)^{s-\sigma}} |w(k)|^2 \leq \\ &\leq \text{Const.} \sum_{|k| > N} \frac{(2\pi \sum_{p=0}^s k^{2p})}{(1 + N^2)^{s-\sigma}} |\hat{w}(k)|^2 = \\ &\leq \text{Const.} \frac{\|w - S_N w\|_{H^s}^2}{N^{2(s-\sigma)}}. \end{aligned}$$

Hence

$$(1.1.28) \quad \|w - S_N w\|_{H^\sigma} \leq \text{Const.}_s \cdot \frac{1}{N^{s-\sigma}}, \quad \sigma \leq s, \quad w \in H^s[0, 2\pi]$$

with $\text{Const.}_s \sim \|w - S_N w\|_{H^s} \xrightarrow{N \rightarrow \infty} 0$. Thus, for each derivative we “lose” one order in the convergence rate.

As a corollary we also get uniform convergence of $S_N w(x)$ for $H^1[0, 2\pi]$ -functions $w(x)$, with the help of Sobolev-type estimate

$$(1.1.29) \quad \text{Max}_{0 \leq x \leq 2\pi} |v(x)| \leq \text{Const.} \|v\|_{H^1}.$$

(Proof: Write $v(x) = \bar{v}(x_0) + \int_{x_0}^x v'(x) dx$ with $\bar{v}(x_0) \equiv \frac{1}{2\pi} \int_0^{2\pi} v(x) dx$, and use Cauchy-Schwartz to upper bound the two integrals on the right.)

Utilizing (1.1.20) with $v(x) = w(x) - S_N w(x)$ we find

$$\begin{aligned} (1.1.30) \quad \text{Max}_{0 \leq x \leq 2\pi} |u(x) - S_N w(x)| &\leq \text{Const.} \|w - S_N w\|_{H^1} \\ &\leq \text{Const.}_s \frac{1}{N^{s-1}} \xrightarrow{N \rightarrow \infty} 0, \quad w \in H^1[0, 2\pi]. \end{aligned}$$

Corollary: Assume $w(x) \in H^1[0, 2\pi]$. Then

$$(1.1.31) \quad w(x) = \sum_{k=-\infty}^{\infty} \hat{w}(k) e^{ikx}.$$

In closing this section, we note that the spectral-Fourier projection, $S_N w(x)$, can be rewritten in the form

$$\begin{aligned}
 (1.1.32a) \quad S_N w(x) &= \sum_{k=-N}^N \hat{w}(k) e^{ikx} = \frac{1}{2\pi} \int_{\xi=0}^{2\pi} w(\xi) \sum_{k=-N}^N e^{ik(x-\xi)} d\xi = \\
 &= \int_{\xi=0}^{2\pi} D_N(x-\xi) w(\xi) d\xi
 \end{aligned}$$

where

$$(1.1.32b) \quad D_N(x-\xi) = \frac{1}{2\pi} \sum_{k=-N}^N e^{ik(x-\xi)} = \frac{1}{2\pi} \frac{\sin\left(N + \frac{1}{2}\right)(x-\xi)}{\sin\left(\frac{x-\xi}{2}\right)}.$$

Thus, the spectral projection is given by a convolution with the so-called Dirichlet kernel,

$$(1.1.33) \quad D_N(x) = \frac{1}{2\pi} \frac{\sin\left(N + \frac{1}{2}\right)x}{\sin\frac{x}{2}}.$$

Now (1.1.23) reads

$$(1.1.34) \quad |w(x) - D_N(x) * w(x)| \leq \text{Const}_s \cdot \frac{1}{N^{s-1}}, \quad \text{Const}_s \sim \|w\|_{H^s}.$$

1.2. The Pseudospectral (Collocation) Approximation

We have seen that given the “moments”

$$\hat{w}(k) = \frac{1}{2\pi} \int_{\xi=0}^{2\pi} w(\xi) e^{-ik\xi} d\xi, \quad -N \leq k \leq N,$$

we can recover smooth functions $w(x)$ within spectral accuracy. Now, suppose we are given discrete data of $w(x)$: specifically, assume $w(x)$ is known at equidistant collocation points

$$(1.2.1) \quad w_\nu = w(x_\nu), \quad x_\nu = r + \nu h, \quad \nu = 0, 1, \dots, 2N.$$

Without loss of generality we can assume that r —which measures a fixed shift from the origin, satisfies

$$(1.2.2) \quad 0 \leq r < h \equiv \frac{2\pi}{2N+1}.$$

Given the equidistant values w_ν , we can approximate the above “moments,” $\hat{w}(k)$, by the trapezoidal rule ²

$$(1.2.3) \quad \tilde{w}(k) = \frac{h}{2\pi} \sum_{\nu=0}^{2N+1} {}''w_\nu e^{-ikx_\nu} \equiv \frac{1}{2N+1} \sum_{\nu=0}^{2N} w_\nu e^{-ikx_\nu}.$$

Using $\tilde{w}(k)$ instead of $\hat{w}(k)$ in (1.1.7), we consider now the pseudospectral approximation

$$(1.2.4) \quad \psi_N w = \sum_{k=-N}^N \tilde{w}(k) e^{ikx}.$$

The error, $w(x) - \psi_N w(x)$, consists of two parts:

$$w(x) - \psi_N w(x) = \sum_{|k|>N} \hat{w}(k) e^{ikx} + \sum_{|k|\leq N} [\hat{w}(k) - \tilde{w}(k)] e^{ikx}.$$

The first contribution on the right is the truncation error

$$(1.2.5) \quad T_N w(x) \equiv (I - S_N) w(x) = \sum_{|k|>N} \hat{w}(k) e^{ikx}.$$

We have seen that it is spectrally small provided $w(x)$ is sufficiently smooth. The second contribution on the right is the aliasing error

$$(1.2.6) \quad A_N w(x) = \sum_{|k|\leq N} [\hat{w}(k) - \tilde{w}(k)] e^{ikx}.$$

This is pure discretization error; to estimate its size we need the

² \sum' and \sum'' indicate summation with $\frac{1}{2}$ of the first and respectively, the first and the last last terms.

Poisson's Summation Formula (Aliasing)

Assume $w(x) \in H^1[0, 2\pi)$. Then we have

$$(1.2.7) \quad \tilde{w}(k) = \sum_{p=-\infty}^{\infty} e^{ip(2N+1)r} \hat{w}(k + p(2N+1)).$$

Proof: For $w(x) \in H^1[0, 2\pi)$ we insert its Fourier expansion

$$(1.2.8) \quad \tilde{w}(k) = \frac{1}{2N+1} \sum_{\nu=0}^{2N} w(x_\nu) e^{-ikx_\nu} = \frac{1}{2N+1} \sum_{\nu=0}^{2N} \left[\sum_{j=-\infty}^{\infty} \hat{w}(j) e^{ijx_\nu} \right] e^{-ikx_\nu}.$$

Since $w(x)$ is assumed to be in H^1 , the summation on the right is absolutely convergent

$$(1.2.9) \quad \sum_{j=-\infty}^{\infty} |\hat{w}(j)| \leq \left(\sum_j (1+j^2) |\hat{w}(j)|^2 \cdot \sum_j \frac{1}{1+j^2} \right)^{\frac{1}{2}} \leq \text{Const.} \|w\|_{H^1},$$

and hence we can interchange the order of summation

$$(1.2.10) \quad \tilde{w}(k) = \frac{1}{2N+1} \sum_{j=-\infty}^{\infty} \hat{w}(j) \sum_{\nu=0}^{2N} e^{i(j-k)x_\nu}.$$

Straightforward calculation yields

$$(1.2.11) \quad \begin{aligned} & \frac{1}{2N+1} \sum_{\nu=0}^{2N} e^{i(j-k)(r+\nu h)} = e^{i(j-k)r} \cdot \frac{1}{2N+1} \cdot \sum_{\nu=0}^{2N} e^{i(j-k)\nu \frac{2\pi}{2N+1}} = \\ & = e^{i(j-k)r} \cdot \frac{1}{2N+1} \begin{cases} \frac{e^{i(j-k)\frac{2\pi(2N+1)}{2N+1}} - 1}{e^{i(j-k)\frac{2\pi}{2N+1}} - 1} = 0 & j-k \neq 0 \pmod{2N+1} \\ 2N+1, & j-k = p \cdot (2N+1). \end{cases} \end{aligned}$$

and we end up with the asserted equality

$$\tilde{w}(k) = \sum_{j=-\infty}^{\infty} \hat{w}(j) \cdot \frac{1}{2N+1} \sum_{\nu=0}^{2N} e^{i(j-k)x_\nu} = \sum_{p=-\infty}^{\infty} \hat{w}(k + p(2N+1)) \cdot e^{ip(2N+1)r}.$$

We note that once $w(x)$ is assumed to be smooth, it is completely determined (pointwise, that is) by its Fourier coefficients $\hat{w}(k)$; so are its equidistant values $w_\nu \equiv w(x_\nu)$ and so are its discrete Fourier coefficients $\tilde{w}(k)$. The last formula shows that $\tilde{w}(k)$ are determined in terms of $\hat{w}(k)$, by folding back high modes on the lowest ones, due to the discrete resolution of the moments of $w(x)$: all modes $= k \pmod{2N+1}$ are aliased at the same place since they are equal on the gridpoints

$$(1.2.12) \quad e^{i(k+p(2N+1))x_\nu} = e^{ip(2N+1)r} \cdot e^{ikx_\nu}.$$

Let us rewrite (1.2.7) in the form

$$\tilde{w}(k) = \hat{w}(k) + \sum_{p \neq 0} e^{ip(2N+1)r} \cdot \hat{w}(k + p(2N+1)).$$

Returning to the aliasing error in (1.26), we now have

$$(1.2.13) \quad A_N w(x) = \sum_{|k| \leq N} \left[\sum_{p \neq 0} e^{ip(2N+1)r} \cdot \hat{w}(k + p(2N+1)) \right] e^{ikx}.$$

We note that $T_N w(x)$ lies outside π_N while $A_N w(x)$ lies in π_N , hence by H^s -orthogonality

$$\begin{aligned} \|w(x) - \psi_N w(x)\|_{H^s}^2 &= \sum_{|k| > N} (1 + |k|^2)^s \cdot |\hat{w}(k)|^2 && \leftarrow \text{truncation} \\ &+ \sum_{|k| \leq N} (1 + |k|^2)^s \cdot \left| \sum_{p \neq 0} e^{ip(2N+1)r} \cdot \hat{w}(k + p(2N+1)) \right|^2 && \leftarrow \text{aliasing.} \end{aligned}$$

Both contributions involve only the high amplitudes – higher than N in absolute value; in fact they involve precisely all of these high amplitudes. This leads us to aliasing estimate

$$\begin{aligned} &\sum_{|k| \leq N} (1 + |k|^2)^s \left| \sum_{p \neq 0} e^{ip(2N+1)r} \cdot \hat{w}(k + p(2N+1)) \right|^2 \leq \\ &\sum_{|k| \leq N} \sum_{p \neq 0} (1 + |k + p(2N+1)|^2)^s |\hat{w}(k + p(2N+1))|^2. \\ (1.2.15) \quad &\cdot \text{Max}_{|k| \leq N} \sum_{p \neq 0} \left[\frac{1 + |k|^2}{1 + |k + p(2N+1)|^2} \right]^s \leq \\ &\|T_N w(x)\|_{H^s}^2 \cdot \sum_{p \neq 0} \left[\frac{1 + N^2}{1 + 4p^2 N^2} \right]^s. \end{aligned}$$

Hence, we conclude that for any $s > \frac{1}{2}$ we have

$$(1.2.16) \quad \|A_N w(x)\|_{H^s} \leq \text{Const}_s \cdot \|T_N w(x)\|_{H^s}, \quad s > \frac{1}{2}.$$

Augmenting this with our previous estimates we end up with spectral accuracy as before, namely

$$(1.2.17) \quad \|w - \psi_N w\|_{H^\sigma} \leq \text{Const}_s \cdot \frac{1}{N^{s-\sigma}}, \quad w \in H^s[0, 2\pi), \quad s \geq \sigma > \frac{1}{2}.$$

We observe that $\psi_N w(x)$ is nothing but the trigonometric interpolant of $w(x)$ at the equidistant points $x = x_\mu$:

$$\begin{aligned} (1.2.18) \quad \psi_N w(x)|_{x=x_\mu} &= \sum_{k=-N}^N \left[\frac{1}{2N+1} \sum_{\nu=0}^{2N} w(x_\nu) e^{-ikx_\nu} \right] e^{ikx_\mu} = \\ &= \sum_{\nu=0}^{2N} w(x_\nu) \cdot \frac{1}{2N+1} \sum_{k=-N}^N e^{ik(\mu-\nu)h} = w(x_\mu). \end{aligned}$$

This shows that ψ_N is in fact a ψ pseudospectral projection, which in the usual sin-cos formulation reads

$$(1.2.19) \quad \psi_N w = \sum_{k=0}^N \tilde{a}_k \cos kx + \tilde{b}_k \sin kx$$

$$\begin{bmatrix} \tilde{a}_k \\ \tilde{b}_k \end{bmatrix} = \frac{2}{2N+1} \sum_{\nu=0}^{2N} w(x_\nu) \begin{bmatrix} \cos kx_\nu \\ \sin kx_\nu \end{bmatrix}.$$

Thus, trigonometric interpolation provides us with an excellent vehicle to perform approximate discretizations with high (= spectral) accuracy, of differential and integral operations. These can be easily carried out in the Fourier space where the exponents serve as eigenfunction. For example, suppose we are given the equidistant gridvalues, w_ν , of an underlying smooth (i.e., also periodic!) function $w(x), w(x) \in H^s[0, 2\pi]$. A second-order accurate discrete derivative is provided by center differencing

$$\frac{dw}{dx}(x = x_\nu) = \frac{w_{\nu+1} - w_{\nu-1}}{2h} + \mathcal{O}(h^2).$$

Note that the error in this case is, $\mathcal{O}(h^2) \equiv w^{(3)}(\xi)h^2$, no matter how smooth $w(x)$ is. Similarly, fourth order approximation is given via Richardson procedure by

$$\frac{dw}{dx}(x = x_\nu) = \frac{8[w_{\nu+1} - w_{\nu-1}] - [w_{\nu+2} - w_{\nu-2}]}{12h} + \mathcal{O}(h^4).$$

The pseudospectral approximation gives us an alternative procedure: construct the trigonometric interpolant

$$(1.2.20) \quad \psi_N w(x) = \sum_{k=-N}^N \tilde{w}(k) e^{ikx}, \quad \tilde{w}(k) = \frac{1}{2N+1} \sum_{\nu=0}^{2N} w_\nu e^{-ikx_\nu}.$$

Differentiate – in the Fourier space this amounts to simple multiplication since the exponentials are eigenfunctions of differentiation,

$$(1.2.21) \quad \frac{d}{dx} \psi_N w(x) = \sum_{k=-N}^N \tilde{w}(k) i k e^{ikx},$$

and we approximate

$$(1.2.22) \quad \frac{dw}{dx}(x = x_\nu) = \frac{d}{dx} \psi_N w(x)|_{x=x_\nu} + \text{spectrally small error}.$$

Indeed, by our estimates we have for $w(x) \in H^s[0, 2\pi], s > 1$,

$$(1.2.23) \quad \max_{0 \leq x \leq 2\pi} \left| \frac{d}{dx} w(x) - \frac{d}{dx} \psi_N w(x) \right| \leq \text{Const.} \|w(x) - \psi_N w(x)\|_{H^2} \leq \frac{\text{Const.}_s}{N^{s-2}}$$

which verifies the asserted spectral accuracy. Similar estimates are valid for higher derivatives. To carry out the above recipe, one proceeds as follows: starting with the vector of gridvalues, $\tilde{w} = (w_0, \dots, w_{2N})$, one computes the discrete Fourier coefficients

$$(1.2.24) \quad \tilde{w}(k) = \frac{1}{2N+1} \sum_{\nu=0}^{2N} w_{\nu} e^{-ikx_{\nu}}, \quad -N \leq k \leq N,$$

or, in matrix formulation

$$(1.2.25) \quad \begin{bmatrix} \tilde{w}(-N) \\ \vdots \\ \tilde{w}(N) \end{bmatrix} = F \begin{bmatrix} w_0 \\ \vdots \\ w_{2N} \end{bmatrix}, \quad F_{k\nu} = \frac{1}{2N+1} e^{-ikx_{\nu}};$$

then we differentiate

$$(1.2.26) \quad \tilde{w}(k) \rightarrow ik\tilde{w}(k),$$

or in matrix formulation

$$(1.2.27) \quad \begin{bmatrix} \tilde{w}(-N) \\ \vdots \\ \tilde{w}(N) \end{bmatrix} \rightarrow \Lambda \begin{bmatrix} \tilde{w}(-N) \\ \vdots \\ \tilde{w}(N) \end{bmatrix}, \quad \Lambda = \begin{bmatrix} -iN & & \\ & \ddots & \\ & & iN \end{bmatrix},$$

and finally, we return to the “physical” space, calculating

$$(1.2.28) \quad \sum_{k=-N}^N ik\tilde{w}(k) e^{ikx_{\nu}}, \quad \nu = 0, 1, \dots, 2N,$$

or in matrix formulation

$$(1.2.29) \quad \begin{bmatrix} \frac{dw}{dx}(x_0) \\ \vdots \\ \frac{dw}{dx}(x_{2N}) \end{bmatrix} = F^* \cdot (2N+1) \begin{bmatrix} -iN\tilde{w}(-N) \\ \vdots \\ iN\tilde{w}(N) \end{bmatrix}, \quad (2N+1)F_{\nu k}^* = e^{ikx_{\nu}}.$$

The summary of these three steps is

$$(1.2.30) \quad \begin{bmatrix} w'(x_0) \\ \vdots \\ w'(x_{2N}) \end{bmatrix} = \psi D \begin{bmatrix} w_0 \\ \vdots \\ w_{2N} \end{bmatrix}, \quad \psi D \equiv (2N+1)F^* \Lambda F,$$

where ψD represents the discrete differentiation matrix, and similarly ψD^s for higher derivatives.

Note: Since $(2N+1)F^*F = I_{2N+1}$ (interpolation!) we apply $\psi D^s = (2N+1)F^* \Lambda^s F$. How does this compare with finite differences and finite-element type differencing?

In periodic second-order differencing we have

$$FD_2 = \frac{1}{2h} \begin{bmatrix} 0 & 1 & \cdots & 0 & -1 \\ -1 & 0 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 0 & 1 \\ 1 & 0 & \cdots & -1 & 0 \end{bmatrix};$$

fourth order differencing yields

$$FD_4 = \frac{1}{12h} \begin{bmatrix} 0 & 8 & -1 & \cdots & 1 & -8 \\ -8 & 0 & & & & 1 \\ 1 & & \ddots & & & \vdots \\ \vdots & & & & & -1 \\ -1 & & & & 0 & 8 \\ 8 & -1 & \cdots & 1 & -8 & 0 \end{bmatrix}.$$

In both cases the second and fourth order differencing takes place in the physical space. The corresponding differencing matrices have finite bandwidth and this reflects the fact that these differencing methods are local. Similarly, finite-element differencing,

$$\frac{1}{6}w'_{\nu-1} + \frac{4}{6}w'_\nu + \frac{1}{6}w'_{\nu+1} = \frac{w_{\nu+1} - w_{\nu-1}}{2h}$$

corresponds to a differencing matrix

$$FE_4 = \begin{bmatrix} \frac{4}{6} & \frac{1}{6} & \cdots & \frac{1}{6} \\ \frac{1}{6} & & \ddots & \frac{1}{6} \\ \frac{1}{6} & \cdots & \frac{1}{6} & \frac{4}{6} \end{bmatrix}^{-1} \cdot \frac{1}{2h} \begin{bmatrix} 0 & 1 & \cdots & 0 & -1 \\ -1 & 0 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 0 & 1 \\ 1 & 0 & \cdots & -1 & 0 \end{bmatrix}.$$

We still operate in physical space with $\mathcal{O}(N)$ operations (tridiagonal solver) and locality is reflected by a very rapid (exponential decay) away from main diagonal. Nevertheless, if we increase the periodic center differences stencil to its limit then we end up with global pseudospectral differentiation

$$(1.2.31) \quad \frac{d}{dx}\psi_N w(x_\nu) = \sum_{k=-N}^N \left(\frac{ik}{2N+1} \sum_{\mu=0}^{2N} w_\mu e^{-ikx_\mu} \right) e^{ikx_\nu};$$

recall the Dirichlet kernel (1.1.33)

$$(1.2.32) \quad \sum_{k=-N}^N e^{ikx} = e^{-iNx} \frac{e^{i(2N+1)x} - 1}{e^{ix} - 1} = \frac{\sin(N + \frac{1}{2})x}{\sin \frac{x}{2}},$$

and its derivative,

$$(1.2.33) \quad \sum_{k=-N}^N i k e^{ikx} = \frac{d}{dx} \frac{\sin(N + \frac{1}{2})x}{\sin \frac{x}{2}} = \frac{(N + \frac{1}{2}) \cos(N + \frac{1}{2})x \sin \frac{x}{2} - \frac{1}{2} \cos \frac{x}{2} \sin(N + \frac{1}{2})x}{\sin^2 \frac{x}{2}}$$

so that

$$(1.2.34) \quad \sum_{k=-N}^N i k e^{ik(\nu-\mu)h} = \frac{(N + \frac{1}{2}) \cos[(N + \frac{1}{2})(\nu - \mu)h]}{\sin(\frac{x_\nu - x_\mu}{2})}.$$

Hence (1.2.31), (1.2.34) give us

$$(1.2.35) \quad w'(x_\nu) \equiv \frac{d}{dx} \psi_N w(x_\nu) = \sum_{\mu=0}^{2N} \frac{1}{2} \frac{(-1)^{\nu-\mu}}{\sin(\frac{x_\nu - x_\mu}{2})} \cdot w_\mu, \quad [\psi D]_{\nu\mu} = \frac{(-1)^{\nu-\mu}}{2 \sin(\frac{x_\nu - x_\mu}{2})}.$$

In this case ψD is a full $(2N + 1) \times (2N + 1)$ matrix whose multiplication requires $\mathcal{O}(N^2)$ operations; however, we can multiply $\psi D[w]$ efficiently using its spectral representation from (1.2.30),

$$\psi D = (2N + 1) F^* \Lambda F.$$

Multiplication by F and F^* can be carried out by FFT which requires only $\mathcal{O}(N \log N)$ operating and hence the total cost here is almost as good as standard “local” methods, and in addition we maintain spectral accuracy.

We have seen how the pseudospectral differentiation works in the physical space. Next, let's examine how the standard finite-difference/element differencing methods operate in the Fourier space. Again, the essential ingredient is that exponentials play the role of eigenfunctions for this type of differencing. To see this, consider for example the usual second order center differencing, $D_2(h)$, for which we have

$$(1.2.36) \quad D_2(h) e^{ikx}|_{x=x_\nu} = \frac{e^{ikx_{\nu+1}} - e^{-ikx_{\nu-1}}}{2h} = \frac{i \sin(kh)}{h} e^{ikx}|_{x=x_\nu},$$

The term $\frac{i \sin(kh)}{h}$ is called the “symbol” of center differencing. By superposition we obtain for arbitrary grid function (represented here by its trigonometric interpolant)

$$(1.2.37) \quad \psi_N w(x) = \sum_{k=-N}^N \tilde{w}(k) e^{ikx}$$

that

$$(1.2.38) \quad \begin{aligned} \frac{w_{\nu+1} - w_{\nu-1}}{2h} &= D_2(h) \psi_N w = \sum_{k=-N}^N \tilde{w}(k) D_2(h) e^{ikx}|_{x=x_\nu} \\ &= \sum_{k=-N}^N \frac{i \sin(kh)}{h} \tilde{w}(k) e^{ikx}|_{x=x_\nu}. \end{aligned}$$

It is second-order accurate differencing since its symbol satisfies

$$(1.2.39) \quad \frac{i \sin(kh)}{h} = ik + \mathcal{O}(k^3 h^2).$$

Note that for the low modes we have $\mathcal{O}(h^2)$ error (the less significant high modes are differenced with $\mathcal{O}(1)$ error but their amplitudes tend rapidly to zero). Thus we have

$$(1.2.40) \quad \begin{aligned} \left\| \frac{d}{dx} \psi_N w - D_2(h) \psi_N w \right\|^2 &= \sum_{k=-N}^N \left| k - \frac{\sin(kh)}{h} \right|^2 |\tilde{w}(k)|^2 \\ &\leq \text{Const.} h^4 \sum_{|k| \leq N} (1 + |k|^2)^3 |\tilde{w}(k)|^2 \leq \text{Const.} h^4 \cdot \|\psi_N w\|_{H^3}^2, \end{aligned}$$

and this estimate should be compared with the usual

$$\left| \frac{d}{dx} w(x_\nu) - \frac{w_{\nu+1} - w_{\nu-1}}{2h} \right| \leq \text{Const.} h^2 \cdot \text{Max}_x |w^{(3)}(x)|.$$

The main difference between these two estimates lies in the fact that (1.2.41) is local, i.e., we need the smoothness of $w(x)$ only in the neighborhood of $x = x_\nu$ and not in the whole interval. The analogue localization in the Fourier space will be dealt later.

Similarly, we have for fourth order differencing the symbol

$$i \frac{1}{3} \left[4 \frac{\sin kh}{h} - \frac{\sin 2kh}{2h} \right] = ik + \mathcal{O}(k^5 h^4).$$

In general, we have difference operators whose matrix representation, D , is of the form

$$(1.2.41) \quad D = [d_{jk}] \quad -N \leq j, k \leq N$$

and it is periodic and antisymmetric (here $[\ell] \equiv \ell \pmod{2N+1}$)

$$(1.2.42) \quad \begin{aligned} \text{(i) periodicity } d_{jk} &= d_{[k-j]} \\ \text{(ii) antisymmetry } d_{jk} &= -d_{kj}. \end{aligned}$$

Matrices satisfying the periodicity property are called circulant, and they all can be diagonalized by the unitary Fourier matrix

$$(1.2.43) \quad D = U^* \Lambda U, \quad U = (2N+1)^{\frac{1}{2}} \cdot F, \quad U^* U = I_{2N+1}.$$

Indeed, with $p - q = \ell$ we have

$$\begin{aligned}
[U^*DU]_{jk} &= \frac{1}{2N+1} \sum_{p,q=-N}^N e^{ijx_p} \cdot d_{[p-q]} e^{-ikx_q} = \\
&= \frac{1}{2N+1} \sum_{\ell,q=-N}^N e^{ij[r+(q+\ell)h]} d_{[\ell]} e^{-ik(r+qh)} \\
(1.2.44) \quad &= \frac{1}{2N+1} \sum_{\ell,q=-N}^N e^{ij\ell h} d_{[\ell]} \cdot \sum_{q=-N}^N e^{-i(k-j) \cdot (r+qh)} \\
&= \begin{cases} 0 & j \neq k, \\ \sum_{\ell=-N}^N e^{ik\ell h} d_{[\ell]} & j = k, \end{cases}
\end{aligned}$$

and using the antisymmetry we end up with symbols λ_k

$$(1.2.45) \quad \Lambda = \text{diag}(\lambda_{-N}, \dots, \lambda_N), \quad \lambda_k = 2i \sum_{\ell=1}^N d_{\ell} \sin(k\ell h).$$

As an example, we obtain for the finite-element differencing system

$$\begin{aligned}
(1.2.46) \quad \lambda_k &= i \frac{\sin kh}{h} \left(\frac{4}{6} + \frac{1}{6} e^{ikh} + \frac{1}{6} e^{-ikh} \right) = \\
&= \frac{6i}{h} \cdot \frac{\sin(kh)}{4 + 2 \cos(kh)} = ik + \mathcal{O}(h^4).
\end{aligned}$$

In general, the symbols are trigonometric polynomials or rational functions in the “dual variable,” kh , which has “exact” representation on the grid in terms of translation operator (polynomials or rational functions), and accuracy is determined by the ability to approximate the exact differentiation symbol ik for $|k| \sim 1$, see Figure 2.

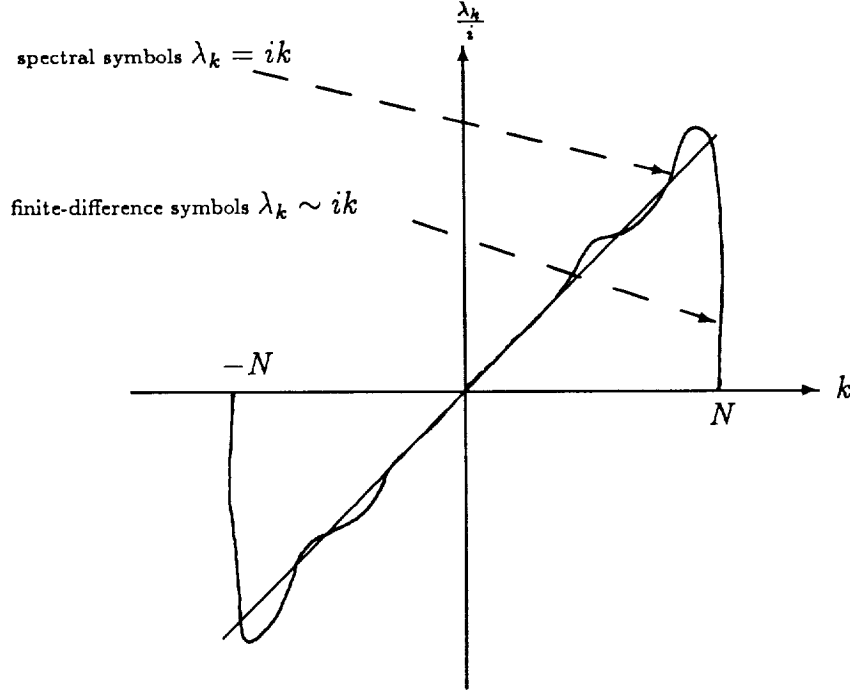


Figure 2:

1.3. Spectral and Pseudospectral Approximations - Exponential Accuracy

We have seen that the spectral and the pseudospectral approximations enjoy what we called “spectral accuracy” – that is, the convergence rate is restricted solely by the global smoothness of the data. The statement about “infinite” order of accuracy for C^∞ functions is an asymptotic statement. Here we show that in the analytic case the error decay rate is in fact exponential.

To this end, assume that

$$(1.3.1) \quad w(z) = \sum_{k=-\infty}^{\infty} \hat{w}(k) e^{ikz}, \quad |Im\ z| \leq \eta < \eta_0,$$

is 2π -periodic analytic in the strip $-\eta_0 < Im\ z < \eta_0$. The error decay rate in both the spectral and pseudospectral cases is determined by the decay rate of the Fourier coefficients $\hat{w}(k)$. Making the change of variables $\zeta = e^{iz}$ we have for

$$(1.3.2) \quad v(\zeta) = w(z = +i\ln\zeta),$$

the power series expansion

$$(1.3.3) \quad v(\zeta) = \sum_{k=-\infty}^{\infty} \hat{w}(k) \zeta^k.$$

By the periodic analyticity of $w(z)$ in the strip $|Imz| \leq \eta < \eta_0$, $v(\zeta)$ is found to be single-valued analytic in the corresponding annulus

$$(1.3.4) \quad e^{-\eta_0} < |\eta| < e^{\eta_0},$$

whose Laurent expansion is given in (1.3.3):

$$(1.3.5) \quad \hat{w}(k) = \frac{1}{2\pi i} \int_{|\zeta|=r} v(\zeta) \zeta^{-(k+1)} d\zeta, \quad e^{-\eta_0} < r < e^{\eta_0}.$$

This yields exponential decay of the Fourier coefficients

$$(1.3.6) \quad |\hat{w}(k)| \leq M(\eta) e^{-k\eta}, \quad M(\eta) = \text{Max}_{|Imz| \leq \eta} |w(z)|, \quad 0 < \eta < \eta_0.$$

We note that the inverse implication is also true; namely an exponential decay like (1.3.6) implies the analyticity of $w(z)$. Inserting this into (1.1.24) yields

$$(1.3.7) \quad \begin{aligned} \|w - S_N w\|^2 &= 2\pi \sum_{|k| > N} |\hat{w}(k)|^2 \leq \\ &\leq 2\pi \cdot M^2(\eta) \cdot \sum_{|k| > N} e^{-2k\eta} = 2\pi \frac{M^2(\eta)}{e^{2\eta} - 1} \cdot e^{-2N\eta} \end{aligned}$$

and similarly for the pseudospectral approximation

$$(1.3.8) \quad \|w - \psi_N w\|^2 \leq \text{Const.} \frac{M^2(\eta)}{e^{2\eta} - 1} \cdot e^{-2N\eta}.$$

Note that in either case the exponential factor depends on the distance of the singularity (lack of analyticity) from the real line. For higher derivatives we likewise obtain

$$(1.3.9) \quad \|w - S_N w\|_{H^\sigma}^2 + \|w - \psi_N w\|_{H^\sigma}^2 \leq \text{Const.} N^{2\sigma} \cdot M^2(\eta) \cdot \frac{e^{-2N\eta}}{e^{2\eta} - 1}.$$

We can do even better, taking into account higher derivatives, e.g.,

$$(1.3.10) \quad k \hat{w}(k) = \frac{1}{2\pi i} \int_{|\zeta|=r} \frac{dv}{d\zeta}(\zeta) \zeta^{-k} d\zeta,$$

so that with

$$(1.3.11) \quad M_s(\eta) = e^{2s\eta} \sum_{j=0}^s \text{Max}_{|\zeta|=e^\eta} |v^{(j)}(\zeta)|,$$

we have

$$(1.3.12) \quad k |\hat{w}(k)| \leq M_1(\eta) e^{-k\eta},$$

and hence

$$(1.3.13) \quad \|w - S_N w\|_{H^\sigma}^2 + \|w - \psi_N w\|_{H^\sigma}^2 \leq \text{Const.} M_\sigma^2(\eta) \frac{e^{-2N\eta}}{e^{2\eta} - 1}.$$

1.4. The Non-Periodic Problem – Chebyshev Approximation

We start by considering the second order SL problem

$$(1.4.1) \quad -\sqrt{1-x^2} \frac{d}{dx} \left(\sqrt{1-x^2} \frac{d}{dx} \psi \right) = \lambda \psi(x), \quad -1 \leq x \leq 1.$$

This is a special case of the general SL problem

$$(1.4.2) \quad L\psi = -\frac{1}{\omega(x)} \frac{d}{dx} \left(p(x) \frac{d\psi}{dx} \right) + \frac{1}{\omega(x)} q(x) \psi(x) = \lambda \psi(x), \quad p, q, w \geq 0.$$

Noting the Green identity

$$(1.4.3) \quad (L\psi, \phi)_{\omega(x)} = \int_a^b -(p\psi')' \phi + q\psi\phi = p(x)[\psi, \phi]_a^b + (\psi, L\phi)_{\omega(x)}, \quad [\psi, \phi] \equiv \psi\phi' - \phi\psi',$$

we find that L is formally self-adjoint provided certain auxiliary conditions are satisfied. In the nonsingular case where $p(a) \cdot p(b) \neq 0$, we augment (1.4.2) with homogeneous boundary conditions,

$$(1.4.4) \quad \psi(a) = \phi(a) = 0, \quad \psi(b) = \phi(b) = 0.$$

Then L is self-adjoint in this case with a complete eigensystem $(\lambda_k, \psi_k(x))$: each $w(x) \in L_{\omega(x)}[a, b]$ has the “generalized” Fourier expansion

$$(1.4.5) \quad w(x) \sim \sum_{k=0}^{\infty} \hat{w}(k) \psi_k(x), \quad \hat{w}(k) = \frac{(w(x), \psi_k(x))_{\omega}}{\|\psi_k(x)\|_{\omega}^2}$$

with Fourier coefficients

$$(1.4.6) \quad \hat{w}(k) = \frac{1}{\|\psi_k\|_{\omega}^2} \int_a^b w(x) \psi_k(x) \omega(x) dx.$$

The decay rate of the coefficients is algebraic: indeed

$$\begin{aligned} \hat{w}(k) &= \frac{1}{\|\psi_k\|_{\omega}^2} \cdot \frac{1}{\lambda_k} (L\psi_k, w)_{\omega} = \\ (1.4.7) \quad &= \frac{1}{\|\psi_k\|_{\omega}^2} \cdot \frac{1}{\lambda_k} (p(x) \cdot [\psi_k, w]_a^b + (\psi_k, Lw)_{\omega}) \\ &= \frac{1}{\|\psi_k\|_{\omega}^2} (p(x) \cdot \sum_{j=0}^{s-1} \frac{1}{\lambda_k^{s+1}} [\psi_k, L^{(j)}w]_a^b + \frac{1}{\lambda_k^s} (\psi_k, L^{(s)}w)_{\omega}). \end{aligned}$$

The asymptotic behavior of the eigenvalues for nonsingular SL problem is

$$\lambda_k \sim \left[\frac{\pi k}{\int_a^b \sqrt{\frac{\omega(x)}{p(x)}} dx} \right]^2 \sim \text{Const.} k^2$$

and hence, unless $w(x)$ satisfies infinite set of boundary restrictions we end with algebraic decay of $\hat{w}(k)$

$$\hat{w}(k) \sim \frac{1}{\|\psi_k\|_\omega^2} \cdot -\frac{p(x)}{\lambda_k} \cdot \psi'_k(x)w(x)|_a^b \sim \frac{\text{Const.}}{k^2}.$$

This leads to algebraic convergence of the corresponding spectral and pseudospectral projections.

In contrast, the singular case is characterized by having, $p(a) = p(b) = 0$, in which case L is self-adjoint independently of the boundary conditions since the brackets $[,]$ drop, and we have spectral decay-compare (1.1.22)

$$(1.4.8) \quad \hat{w}(k) = \frac{1}{\|\psi_k\|_\omega^2} \cdot \frac{1}{\lambda_k^s} \cdot (\psi_k, L^{(s)}w)_\omega \leq \frac{1}{\lambda_k^s} \frac{\|L^{(s)}w\|_\omega}{\|\psi_k\|_\omega},$$

provided $w(x)$ is smooth enough; that is, the decay is as rapid as the smoothness of $w(x)$ may permit.

Returning to the singular SL problem (1.4.1) we use the transformation

$$(1.4.9) \quad x = \cos \theta, \quad \frac{d}{dx} = \frac{1}{\frac{dx}{d\theta}} \cdot \frac{d}{d\theta} = -\frac{1}{\sqrt{1-x^2}} \frac{d}{d\theta}$$

which yields

$$(1.4.10) \quad -\frac{d^2}{d\theta^2} \phi(\theta) = \lambda \phi(\theta), \quad \phi(\theta) \equiv \psi(\cos \theta),$$

obtaining the two sets of eigensystems

$$(1.4.11a) \quad (\lambda_k = k^2, \phi_k = \cos k\theta),$$

and

$$(1.4.11b) \quad (\lambda_k = k^2, \phi_k = \sin k\theta).$$

The second set violates the boundedness requirement which we now impose

$$(1.4.12) \quad |\psi'_k(\pm 1)| \leq \text{Const.},$$

and so we are left with

$$(1.4.13) \quad (\lambda_k = k^2, \psi_k(x) = \cos(k \cos^{-1} x)).$$

The trigonometric identity

$$\cos(k+1)\theta = 2 \cos \theta \cos k\theta - \cos(k-1)\theta$$

yields the recurrence relation

$$(1.4.14) \quad \psi_{k+1}(x) = 2x\psi_k(x) - \psi_{k-1}(x), \quad \psi_0(x) \equiv 1, \psi_1(x) = x,$$

hence, $\psi_k(x)$ are polynomials of degree k – these are the Chebyshev polynomials

$$(1.4.15) \quad T_k(x) = \cos(k \cos^{-1} x)$$

which are orthonormal w.r.t. Chebyshev weight $\omega(x) = (1 - x^2)^{-\frac{1}{2}}$,

$$(1.4.16) \quad (T_k(x), T_j(x))_\omega = \int_{-1}^1 \frac{T_k(x)T_j(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 & j \neq k, \\ \|T_k\|_\omega^2 = \frac{\pi}{2} & j = k > 0, \\ \|T_0\|_\omega^2 = \pi & j = k = 0. \end{cases}$$

In analogy with what we had done before, we consider now the Chebyshev-Fourier expansion

$$(1.4.17) \quad w(x) \sim \sum_{k=0}^{\infty} \hat{w}(k) T_k(x), \quad \hat{w}(k) = \frac{(w(x), T_k(x))_\omega}{\|T_k\|_\omega^2}.$$

To get rid of the factor $\frac{1}{2}$ for $k = 0$ we may also write this as

$$(1.4.18a) \quad w(x) \sim \sum_{k=0}^{\infty} ' \hat{w}(k) T_k(x)$$

where

$$(1.4.18b) \quad \hat{w}(k) = \frac{(w(x), T_k(x))_\omega}{\pi/2} = \frac{2}{\pi} \int_{-1}^1 \frac{w(x) \cos(k \cos^{-1} x) dx}{\sqrt{1-x^2}}$$

or using the above Chebyshev transformation

$$(1.4.18c) \quad \hat{w}(k) = \frac{2}{\pi} \int_{\xi=0}^{\pi} w(\cos \xi) \cos k\xi \, d\xi.$$

Thus, we go from the interval $[-1, 1]$ into the 2π -periodic circle by even extension, with Fourier expansion of $w(\cos \theta)$, compare (1.1.9) and see Figure 3,

$$\hat{w}(k) = \frac{1}{\pi} \int_{\xi=0}^{2\pi} w(\cos \xi) \cos k\xi \, d\xi = \frac{2}{\pi} \int_{\xi=0}^{\pi} w(\cos \xi) \cos k\xi \, d\xi.$$

Another way of writing this employs a symmetric doubly infinite Fourier-like summation, where

$$(1.4.19) \quad w(x) \sim \frac{1}{2} \sum_{k=-\infty}^{\infty} \hat{w}(k) T_k(x)$$

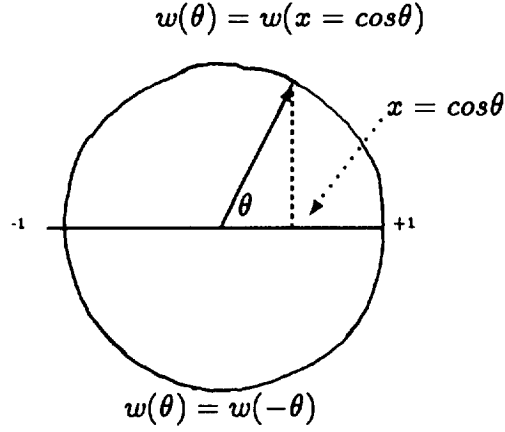


Figure 3:

with $T_{-k}(x) \equiv T_k(x)$ and

$$(1.4.20) \quad \hat{w}(k) = \frac{2}{\pi} \int_{-1}^1 \frac{w(x)T_k(x)}{\sqrt{1-x^2}} dx, \quad -\infty < k < \infty.$$

The Parseval identity reflects the completeness of this system

$$(1.4.21) \quad \begin{aligned} \|w(x)\|_T^2 &\equiv \int \frac{w^2(x)dx}{\sqrt{1-x^2}} = \frac{1}{4} \left[\pi |\hat{w}(0)|^2 + \frac{\pi}{2} \sum_{k \neq 0} |\hat{w}(k)|^2 \right] \\ &= \frac{\pi}{4} \sum_{k=0}^{\infty} |\hat{w}(k)|^2 \end{aligned}$$

which yields the error estimate

$$\|w - S_N w\|_T^2 = \frac{\pi}{4} \sum_{k > N} |\hat{w}(k)|^2.$$

Now, in order to get a measure of the spectral convergence we have to estimate the decay rate of Chebyshev coefficients in terms of the smoothness of $w(x)$ and its derivatives; to this end we need Sobolev like norms. Unlike the Fourier case, $\{T_k(x)\}$ is not complete with respect to H^s – orthogonality is lost because of the Chebyshev weight. So we can proceed formally as before, see (1.1.24),

$$(1.4.22) \quad \|w - S_N w\|_T^2 = 2\pi \sum_{k > N} |\hat{w}(k)|^2 \leq \sum_{k > N} \frac{(1 + |k|^2)^s}{(1 + N^2)^s} |\hat{w}(k)|^2$$

i.e., if we define the Chebyshev-Sobolev norm

$$\|w\|_{H_T^s}^2 = \sum_{k=0}^{\infty} (1 + |k|^2)^s |\hat{w}(k)|^2,$$

then we have spectral accuracy

$$\|w - S_N w\|_T \leq \text{Const}_s \cdot \frac{1}{N^s}, \quad w \in H_T^s[-1, 1].$$

In fact the H_T^s space can be derived from appropriate inner product in the real space as done in Fourier expansion. The correct inner product is given by – compare (1.1.19)

$$(1.4.23) \quad (w_1, w_2)_{H_T^{2s}} = \sum_{p=0}^s (L^p w_1, L^p w_2)_T \underset{x=\cos\theta}{=} \sum_{p=0}^s \int_{\theta=0}^{2\pi} \frac{d^{2p}}{d\theta^{2p}} w_1(\cos\theta) \frac{d^{2p}}{d\theta^{2p}} w_2(\cos\theta) d\theta,$$

so that

$$(1.4.24) \quad (T_k, T_j)_{H_T^{2s}} = \begin{cases} 0 & j \neq k, \\ \frac{\pi}{2} \sum_{p=0}^s k^{4p}, & j = k \text{ (with } \pi \text{ factor at } j = k = 0). \end{cases}$$

Hence the Fourier coefficients in this Hilbert space behave like

$$(1.4.25) \quad (w(x), T_k)_{H_T^{2s}} \sim \sum_{k=0}^{\infty} (1 + k^2)^{2s} \hat{w}(k),$$

and the corresponding norm is equivalent to

$$(1.4.26) \quad \|w\|_{H_T^{2s}}^2 \sim \sum_{k=0}^{\infty} (1 + k^2)^{2s} |\hat{w}(k)|^2.$$

The reason for the squared factors here is due to the fact that L is a second order differential operator, unlike the first-order $D = \frac{d}{dx}$ in the Fourier case, i.e.,

$$(1.4.27) \quad \sum_{k=0}^{\infty} (1 + |k|^2)^{2s} |\hat{w}(k)|^2 \sim \sum_{p=0}^s \|L^p w\|_T^2$$

involves the first $2s$ -derivatives of $w(x)$ – appropriately weighted by Chebyshev weight. This completes the analogy with the Fourier case, and enables us to estimate derivative as well – compare (1.1.28 - 1.1.29),

$$(1.4.28) \quad \|w - S_N w\|_{H_T^{\sigma}} \leq \text{Const}_s \frac{1}{N^{s-\sigma}}, \quad \sigma \leq s, \quad w \in H_T^s[-1, 1].$$

Next, let's discuss the discrete setup. Because we seek an even extension of the upper semi-circle we consider the case of even number of grid points – equally distributed along the unit circle. One choice is to look at

$$\theta_{\nu} = r + \nu h = \left(\nu + \frac{1}{2}\right) \frac{\pi}{N} \quad \left(h = \frac{\pi}{N}, r = \frac{h}{2}\right)$$

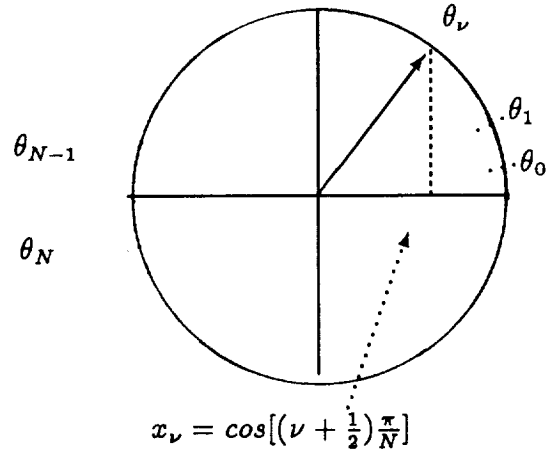


Figure 4:

which gives the usual Chebyshev interpolation points, see Figure 4. The second choice – which we will concentrate from now on – takes into account the boundaries, where we look at

$$\theta_\nu = \nu h = \nu \frac{\pi}{N} \quad (h = \frac{2\pi}{2N}, \quad r = 0)$$

which yields, see Figure 5,

$$(1.4.29) \quad x_\nu = \cos(\nu h), \quad \nu = 0, 1, \dots, N.$$

The trapezoidal rule is replaced by Gauss-Lobatto rule for (1.4.20), i.e., starting with

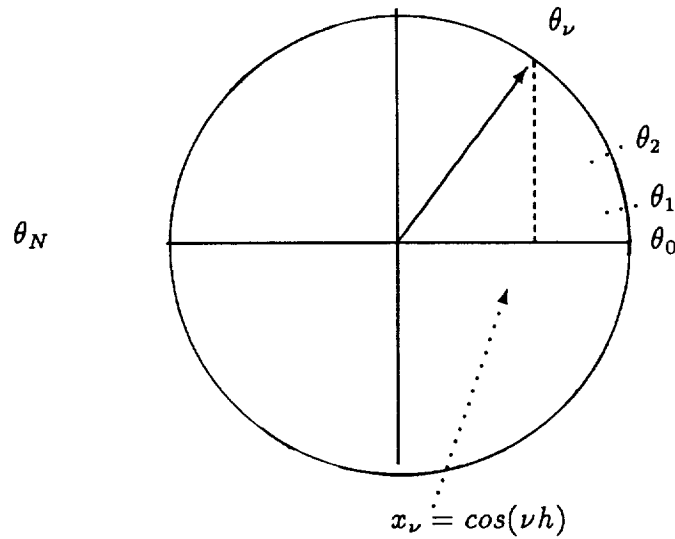


Figure 5:

(1.4.18c), we have

$$(1.4.30) \quad \hat{w}(k) = \frac{2}{\pi} \int_{\xi=0}^{\pi} w(\cos \xi) \cos k\xi d\xi \rightarrow \frac{2}{\pi} \sum_{\nu=0}^N {}''w(\cos \theta_{\nu}) \cos k\theta_{\nu} \cdot \frac{\pi}{N},$$

and we end up with the discrete Chebyshev coefficients

$$(1.4.31) \quad \tilde{w}(k) = \frac{2}{N} \sum_{\nu=0}^N {}''w_{\nu} T_k(x_{\nu}), \quad 0 \leq k \leq N.$$

Indeed, this corresponds to the case of Fourier interpolation with even number of equidistant gridpoints θ_{ν} in (A.1.2), for

$$\begin{aligned} \tilde{w}(k) &= \frac{1}{2\pi} \sum_{\nu=0}^{2N} {}''w_{\nu} e^{-ik\theta_{\nu}} h = \frac{1}{\pi} \sum_{\nu=0}^N w_{\nu} [e^{-ik\theta_{\nu}} + e^{ik\theta_{\nu}}] \frac{2\pi}{2N} = \\ &= \frac{2}{N} \sum_{\nu=0}^N {}''w_{\nu} \cos(k\theta_{\nu}). \end{aligned}$$

Then one may construct the Chebyshev interpolant at these $N + 1$ gridpoints

$$(1.4.32) \quad \psi_N w(x) = \sum_{k=0}^N {}''\tilde{w}(k) T_k(x).$$

We have an identical aliasing relation

$$(1.4.33) \quad \tilde{w}(k) = \sum_{p=-\infty}^{\infty} \hat{w}(k + 2pN)$$

(compare A.1.5), and hence spectral convergence, i.e., (compare (1.2.14) and (1.2.12))

$$(1.4.34) \quad \|w(x) - \psi_N w(x)\|_{H_T^s} \leq \text{Const}_s \cdot \frac{1}{N^{s-\sigma}}, \quad w \in H_T^s, \quad s \geq \sigma,$$

where $\text{Const}_s \sim \|w\|_{H_T^s}$.

Example: We have the Sobolev imbedding of L^{∞} space

$$\begin{aligned} |w(x)| &\leq \frac{1}{2} \sum_{k=-\infty}^{\infty} |\hat{w}(k)| \leq \frac{1}{2} \left(\sum_k (1 + k^2)^{\sigma} |\hat{w}(k)|^2 \cdot \sum_k \frac{1}{(1 + k^2)^{\sigma}} \right)^{\frac{1}{2}} \\ &\leq \text{Const}_{\sigma} \cdot \|w\|_{H_T^{\sigma}}, \quad \sigma > \frac{1}{2}. \end{aligned}$$

Consequently,

$$\text{Max}_x |w(x) - \psi_N w(x)| \leq (\text{Const}_s \sim \|w\|_{H_T^s}) \cdot \frac{1}{N^{s-\sigma}}, \quad s \geq \sigma > \frac{1}{2}.$$

In particular, with $s = N + 1$ we obtain the "usual" estimate for the near minmax approximation (collocated at $x_\nu = \cos\left(\left(\nu + \frac{1}{2}\right)\frac{\pi}{N}\right)$)

$$\text{Max}_x |w(x) - \psi_N w(x)| \leq \text{Const.} \|w\|_{H_T^{N+1}} \cdot \frac{e^{-N}}{N!}.$$

We briefly mention the exponential convergence in the analytic case. To this end we employ Bernstein's regularity ellipse, E_r , with foci ± 1 and sum of its semi axis $= r$. Denoting

$$(1.4.35) \quad M(\eta) = \text{Max}_{z \in E_r} |w(z)|, \quad r = e^\eta.$$

We have

Theorem. Assume $w(x)$ is analytic in $[-1, 1]$ with regularity ellipse whose sum of semiaxis $= r_0 = e^{\eta_0} > 1$. Then

$$\|w(x) - \psi_N w(x)\|_{H^\sigma}^2 + \|w(x) - S_N w(x)\|_{H^\sigma}^2 \leq \text{Const.} \frac{M^2(\eta)}{e^{2\eta} - 1} \cdot N^{2\sigma} e^{-2N\eta}.$$

Proof: The transformation $z = (\zeta + \zeta^{-1})/2$ takes E_{r_0} from the z -plane into the annulus $r_0^{-1} < |\zeta| < r_0$ in the ζ -plane. Hence, $v(\zeta) = 2w\left(z = \frac{\zeta + \zeta^{-1}}{2}\right)$ admits the power expansion

$$(1.4.36) \quad v(\zeta) = 2w\left(\frac{\zeta + \zeta^{-1}}{2}\right) = \sum_{k=-\infty}^{\infty} \hat{w}(k) \zeta^k, \quad r_0^{-1} < |\zeta| < r_0 = e^{\zeta_0};$$

indeed, setting $\zeta = e^{i\theta}$ and recalling $\hat{w}(-k) = \hat{w}(k)$, the above expansion clearly describes the real interval $[-1, 1]$

$$(1.4.37) \quad w(z = \cos \theta) = \sum_{k=0}^{\infty} \hat{w}(k) \cos k\theta.$$

By Laurent expansion in (1.4.36)

$$(1.4.38) \quad \hat{w}(k) = \frac{1}{2\pi i} \int_{|\zeta|=r} \frac{\nu(\zeta)}{\zeta^{k+1}} d\zeta, \quad e^{-\eta_0} < r < e^{\eta_0},$$

hence

$$(1.4.39) \quad |\hat{w}(k)| \leq M(\eta) e^{-k\eta}$$

and the result follows along the lines of (1.3.7-8).

Finally, we conclude with discussion on Chebyshev differencing. Starting with grid values w_ν at Chebyshev points $x_\nu = \cos\left(\nu \frac{\pi}{N}\right)$, one constructs the Chebyshev interpolant

$$(1.4.40) \quad \psi_N w(x) = \sum_{k=0}^N \tilde{w}(k) \cdot T_k(x), \quad \tilde{w}(k) = \frac{2}{N} \sum_{\nu=0}^N w_\nu \cos(kx_\nu).$$

One can compute $\tilde{w}(k)$, $0 \leq k \leq N$, efficiently via the cos-FFT with $\mathcal{O}(N \log N)$ operations. Next, we differentiate in Chebyshev space

$$(1.4.41) \quad \frac{d}{dx} \psi_N w(x) = \sum_{k=0}^N {}''\tilde{w}(k) \frac{d}{dx} T_k(x).$$

In this case, however, $T_k(x)$ is not an eigenfunction of $\frac{d}{dx}$; instead $\frac{d}{dx} T_k(x)$ – being a polynomial of degree $\leq k-1$, can be expressed as linear combination of $\{T_j(x)\}_{j=0}^{k-1}$ (in fact $T_k(x)$ is even/odd for even/odd k 's): with $c_0 = 2, c_{k>0} = 1$ we obtain

$$(1.4.42) \quad \frac{d}{dx} T_k(x) = \frac{2}{c_k} \sum_{\substack{0 \leq j < k \\ k-j \text{ odd}}} T_j(x),$$

and hence

$$(1.4.43) \quad \frac{d}{dx} \psi_N w(x) = \sum_{k=0}^N \tilde{w}(k) \frac{2}{c_k} \sum_{\substack{0 \leq j < k \\ k-j \text{ odd}}} T_j(x).$$

Rearranging we get

$$(1.4.44) \quad \frac{d}{dx} \psi_N w(x) = \sum_{k=0}^N {}''\tilde{w}'(k) T_k(x), \quad \tilde{w}'(k) = \frac{2}{c_k} \sum_{\substack{p \geq k+1 \\ p+k \text{ odd}}} p \tilde{w}(p)$$

and similarly for the second derivative

$$(1.4.45) \quad \tilde{w}''(k) = \frac{2}{c_k} \sum_{\substack{p \geq k+2 \\ p+k \text{ even}}} p(p^2 - k^2) \tilde{w}(p).$$

The amount of work to carry out the differentiation in this form is $\mathcal{O}(N^2)$ operations which destroys the $N \log N$ efficiency. Instead, we can employ the recursion relation which follows directly from (1.4.44)

$$(1.4.46) \quad \tilde{w}'(k+1) = \tilde{w}'(k-1) \cdot c_{k-1} - 2k \tilde{w}(k).$$

To see this in a different way we note that

$$\sin(k+1)\theta = \sin(k-1)\theta + 2 \sin \theta \cos k\theta,$$

which leads to

$$\frac{1}{k+1} \frac{dT_{k+1}}{dx} = \frac{1}{k-1} \frac{dT_{k-1}}{dx} + 2T_k(x),$$

and hence

$$\begin{aligned} \frac{d}{dx} \psi_N w(x) &= \sum_{k=0}^N {}''k \tilde{w}(k) \frac{1}{k} T_k'(x) = \\ &= \frac{1}{2} \sum_{k=0}^N {}''(\tilde{w}'(k-1) - \tilde{w}'(k+1)) \frac{1}{k} T_k'(x) \leftarrow \text{summation by parts} \\ &= \frac{1}{2} \sum_{k=0}^N {}''2 \tilde{w}'(k) T_k(x) = \sum_{k=0}^N {}''\tilde{w}'(k) T_k(x) \end{aligned}$$

as asserted. In general we have

$$(1.4.47) \quad \tilde{w}^{(s)}(k+1) = \tilde{w}^{(s)}(k-1)c_{k-1} - 2k\tilde{w}^{(s-1)}(k).$$

With this $\tilde{w}(k)$ can be evaluated using $\mathcal{O}(N)$ operations, and the differentiated polynomial at the grid points is computed using another cos-FFT employing $\mathcal{O}(N \log N)$ operations

$$(1.4.48) \quad \frac{d}{dx}\psi_N w(x)|_{x=x_\nu} = \sum_{k=0}^N {}''\tilde{w}'(k) \cos kx_\nu,$$

with spectral/exponential error

$$(1.4.49) \quad \text{Max}_{x=x_\nu} \left| \frac{d}{dx}w(x) - \frac{d}{dx}\psi_N w(x) \right| \leq \begin{cases} \text{Const}_s \cdot \frac{1}{N^{s-\sigma}} & \frac{3}{2} < \sigma < s, \\ \text{Const}_s \cdot e^{-N\eta} & \end{cases}.$$

The matrix representation of Chebyshev differentiation, D_T , takes the almost antisymmetric form

$$(D_T)_{jk} = \begin{cases} \frac{c_j (-1)^{j+k}}{c_k x_j - x_k} & j \neq k, \\ -\frac{x_j}{2(1-x_j^2)} & j = k \neq (0, N), \\ \frac{2N^2+1}{6} & j = k = 0, \\ -\frac{2N^2+1}{6} & j = k = N. \end{cases}$$

2. TIME DEPENDENT PROBLEMS

2.1. Initial Value Problems of Hyperbolic Type

The wave equation,

$$(2.1.1) \quad w_{tt} = a^2 w_{xx},$$

is the prototype of P.D.E.'s of hyperbolic type. We study the pure initial-value problem associated with (2.1.1), augmented with 2π -periodic boundary conditions and subject to prescribed initial conditions,

$$(2.1.2) \quad w(x, 0) = f(x), \quad w_t(x, 0) = g(x).$$

We can solve this equation using the method of characteristics, which yields

$$(2.1.3) \quad w(x, t) = \frac{f(x + at) + f(x - at)}{2} + \frac{1}{2a} \int_{x-at}^{x+at} g(s) ds.$$

We shall study the manner in which the solution depends on the initial data. In this context the following features are of importance.

1. Linearity: the principle of superposition holds.
2. Finite speed of propagation: influence propagates with speed $\leq a$. This is the essential feature of hyperbolicity. In the wave equation it is reflected by the fact that the value of w at (x, t) is not influenced by initial values outside domain of dependence $(x - at, x + at)$.
3. Existence for large enough set of admissible initial data: arbitrary C_0^∞ initial data can be prescribed and the corresponding solution is C_0^∞ .
4. Uniqueness: the solution is uniquely determined for $-\infty < t < \infty$ by its initial data.
5. Conservation of Energy. The wave equation (2.1.1) describes the motion of a string with kinetic energy, $\frac{1}{2}\rho \int w_t^2 dx$, and potential one, $\frac{1}{2}T \int w_x^2 dx$, ($T/\rho = a^2$). In order to show that the total energy

$$E_{\text{Total}} = \frac{1}{2}\rho \int (w_t^2 + a^2 w_x^2) dx,$$

is conserved in time we may proceed in one of two ways: either by the so called energy method or by the Fourier method.

The Energy Method.

Rewrite (2.1.1) as a first order system

$$(2.1.4a) \quad \frac{\partial}{\partial t} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 & a^2 \\ 1 & 0 \end{bmatrix} \frac{\partial}{\partial x} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial w}{\partial t} \\ \frac{\partial w}{\partial x} \end{bmatrix},$$

or equivalently,

$$(2.1.4b) \quad \frac{\partial u}{\partial t} = A \frac{\partial u}{\partial x}.$$

The essential ingredient here is the existence of a positive symmetrizer, $H > 0$,

$$(2.1.5) \quad HA = \begin{bmatrix} 0 & a^2 \\ a^2 & 0 \end{bmatrix} \equiv A_s = A_s^T, \quad H = \begin{bmatrix} 1 & 0 \\ 0 & a^2 \end{bmatrix},$$

so that multiplication by H on the left gives

$$(2.1.6) \quad Hu_t = A_s u_x.$$

Now, multiplying by u^T we are led to

$$(2.1.7) \quad (u, Hu_t) = (u, A_s u_x),$$

and the real part of both sides are in fact perfect derivatives, for by the symmetry of H ,

$$\begin{aligned} \operatorname{Re}(u, Hu_t) &= \frac{1}{2}(u, Hu_t) + \frac{1}{2}(Hu_t, u) = \\ &= \frac{1}{2}(u, Hu_t) + \frac{1}{2}(u_t, Hu) = \frac{\partial}{\partial t} \left[\frac{1}{2}(u, Hu) \right], \end{aligned}$$

and similarly, by the symmetry of A_s , we have

$$\operatorname{Re}(u, A_s u_x) = \frac{1}{2}(u, A_s u_x) + \frac{1}{2}(A_s u_x, u) = \frac{\partial}{\partial x} \left[\frac{1}{2}(u, A_s u) \right].$$

Hence, by integration over the 2π -period we end up with energy conservation, asserting

$$(2.1.8) \quad \frac{d}{dt} \int_x (w_t^2 + a^2 w_x^2) dx = \frac{d}{dt} \int_x (u, Hu) dx = \int_x \frac{\partial}{\partial x} (u, A_s u) dx = 0.$$

We note that the positivity of H was not used in the proof and is assumed just for the sake of making (u, Hu) an admissible convex “energy norm.”

The Fourier Method.

Fourier transform (2.1.4b) to get the O.D.E.

$$(2.1.9) \quad \frac{\partial \hat{u}}{\partial t}(k, t) = ikA\hat{u}(k, t),$$

whose solution is

$$(2.1.10) \quad \hat{u}(k, t) = e^{ikAt}\hat{u}(k, 0),$$

where $\hat{u}(k, 0)$ is the Fourier transform of the initial data. Now, for

$$(2.1.11) \quad A = T\Lambda T^{-1}, \quad \Lambda = \begin{bmatrix} -a & \\ & a \end{bmatrix}, T = \begin{bmatrix} -a & a \\ 1 & 1 \end{bmatrix},$$

we find

$$(2.1.12) \quad \hat{u}(k, t) = Te^{ik\Lambda t}T^{-1}\hat{u}(k, 0)$$

or

$$(2.1.13) \quad T^{-1}\hat{u}(k, t) = \begin{bmatrix} e^{-ikat} & 0 \\ 0 & e^{ikat} \end{bmatrix} T^{-1}\hat{u}(k, 0)$$

and hence (since the diagonal matrix on the right is clearly unitary), the L_2 -norm of $T^{-1}\hat{u}(k, t)$ is conserved in time, i.e.,

$$(2.1.14) \quad \|T^{-1}\hat{u}(k, t)\|^2 = \|T^{-1}\hat{u}(k, 0)\|^2, \quad T^{-1} = -\frac{1}{2a} \begin{bmatrix} 1 & -a \\ -1 & -a \end{bmatrix}.$$

Summing over all modes and using Parseval we end up with energy conservation

$$(2.1.15a) \quad \begin{aligned} \int_x (w_t^2 + a^2 w_x^2) dx &= 4a^2 \int_x \left(\frac{w_t - aw_x}{-2a} \right)^2 + \left(\frac{w_t + aw_x}{-2a} \right)^2 dx \\ &= 4a^2 \int_x \|T^{-1}u\|^2 dx = 8\pi a^2 \sum_k \|T^{-1}\hat{u}(k, t)\|^2 = \text{Const.} \end{aligned}$$

as asserted.

We note that the only tool used in the energy method was the existence of a positive symmetrizer for A , while the only tool used in the Fourier method was the real diagonalization of A ; in fact the two are related, for if $A = T\Lambda T^{-1}$ then for $H = (T^{-1})^* T^{-1} > 0$ we have

$$(2.1.15b) \quad HA = (T^{-1})^* \Lambda T^{-1} = A_\bullet \equiv A_\bullet^T, \quad \Lambda \text{ real diagonal.}$$

Energy conservation implies (in view of linearity) uniqueness, and serves as a basic tool to prove existence. It will be taken as definition of hyperbolicity. It implies and is implied by the qualitative properties (1) - (4).

We now turn to consider general P.D.E.'s of the form

$$(2.1.16a) \quad \frac{\partial u}{\partial t} = P(x, t, D)u, \quad P(x, t, D) = \sum_{j=1}^d A_j(x, t) \frac{\partial}{\partial x_j},$$

with 2π -periodic boundary conditions and subject to prescribed initial conditions,

$$(2.1.16b) \quad u(x, 0) = f(x).$$

We say that the system (2.1.16a) is hyperbolic if the following a priori energy estimate holds:

$$(2.1.17) \quad \|u(x, t)\|_{L_2(x)} \leq \text{Const.}_T \cdot \|u(x, 0)\|_{L_2(x)}, \quad -T \leq t \leq T.$$

As we shall see later on this is equivalent to energy conservation with appropriate energy renorming.

Here are the basic facts concerning such systems.

The Constant Coefficients Case.

$$(2.1.18) \quad \frac{\partial u}{\partial t} = P(D)u, \quad P(D) = \sum_{j=1}^d A_j \frac{\partial}{\partial x_j}, \quad A_j = \text{constant matrices}.$$

Define the Fourier symbol associated with $P(D)$:

$$(2.1.19) \quad \hat{P}(ik) = i \sum_{j=1}^d A_j k_j, \quad k = (k_1, k_2, \dots, k_d) \in R^d,$$

which naturally arises as we Fourier transform (2.1.18),

$$(2.1.20) \quad \frac{\partial}{\partial t} \hat{u}(k, t) = \hat{P}(ik) \hat{u}(k, t).$$

Solving the O.D.E. (2.1.20) we find, as before, that hyperbolicity amounts to

$$(2.2.21) \quad \|e^{\hat{P}(ik)t}\| \leq \text{Const}_T, \quad -T \leq t \leq T, \quad \text{for all } k' \text{'s}.$$

For this to be true the necessary Garding-Petrovski condition should hold, namely

$$(2.1.22) \quad |\text{Re} \lambda[\hat{P}(ik)]| \leq \text{Const}.$$

Example: For the wave equation, $\lambda[\hat{P}(ik)] = \pm ika$.

But the Garding-Petrovski condition is not sufficient for the hyperbolic estimate (2.1.17) as told by the counterexample

$$\frac{\partial}{\partial t} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix} \frac{\partial}{\partial x} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

As before, in this case we have $\lambda[\hat{P}(ik)] = \pm ika$, hence the Garding-Petrovski condition is fulfilled. Yet, Fourier analysis shows that we need both $\|u_1(x, 0)\|_{L_2(x)}$ and $\|\frac{\partial u_2}{\partial x}(x, 0)\|_{L_2(x)}$ in order to upperbound $\|u_1(x, t)\|_{L_2(x)}$. Thus, the best we can hope for with this counterexample is an a priori estimate of the form

$$\|u(x, t)\|_{L_2(x)} \leq \text{Const}_T \cdot \|u(x, 0)\|_{H^1(x)}, \quad -T \leq t \leq T.$$

We note that in this case we have a "loss" of one derivative. This brings us to the notion of weak hyperbolicity.

We say that the system (2.1.16a) is weakly hyperbolic if there exists an $s \geq 0$ such that the following a priori estimate holds:

$$\|u(x, t)\|_{L_2(x)} \leq \text{Const}_T \cdot \|u(x, 0)\|_{H^s(x)}, \quad -T \leq t \leq T.$$

The Garding-Petrovski condition is necessary and sufficient for the system (2.1.18) to be weakly hyperbolic. The necessary and sufficient characterization of hyperbolic systems is provided by the Kreiss matrix theorem. It states that (2.1.21) holds iff there exists a positive symmetrizer $\hat{H}(k)$ such that

$$(2.1.23) \quad \text{Re}[\hat{H}(k)\hat{P}(ik)] \equiv 0, \quad 0 < m \leq \hat{H}(k) \leq M,$$

and this yields the conservation for $\|u(x, t)\|_H^2 = 2\pi \sum_k \|\hat{u}(k)\|_{\hat{H}(k)}^2$, i.e.,

$$2\pi \sum_k (\hat{u}(k, t), \hat{H}(k)\hat{u}(k, t))$$

is conserved in time.

Remark: For an a priori estimate forward in time ($0 \leq t \leq T$), it will suffice to have

$$(2.1.24) \quad \text{Re}[\hat{H}(k)\hat{P}(ik)] = \frac{1}{2}[\hat{H}(k)\hat{P}(ik) + \hat{P}(ik)\hat{H}(k)] \leq 0.$$

Indeed, we have in this case

$$\frac{1}{2} \frac{d}{dt} (\hat{u}(k), \hat{H}(k)\hat{u}(k)) \leq (\text{Re}[\hat{H}(k)\hat{P}(ik)]\hat{u}(k), \hat{u}(k)) \leq 0,$$

and hence summing over all k 's and using Parseval's

$$\|u(x, t)\|_{L_2(x)}^2 \leq \frac{M}{m} \cdot \|u(x, 0)\|_{L_2(x)}^2.$$

Special important cases are the strictly hyperbolic systems where $\hat{P}(ik)$ has distinct real eigenvalues, so that $\hat{P}(ik)$ can be real diagonalized

$$\hat{P}(k) = iT(k)\Lambda(k)T^{-1}(k),$$

and, as before, $\hat{H}(k) = (T^{-1}(k))^* T^{-1}(k)$ will do. The other important case consists of symmetric hyperbolic systems which can be symmetrized in the physical space, i.e. there exists an $H > 0$ such that

$$HA_j = A_{j*} = A_{j*}^T.$$

Most of the physically relevant systems fall into these categories.

Example: Shallow water equations (linearized)

$$\frac{\partial}{\partial t} \begin{bmatrix} u \\ v \\ \phi \end{bmatrix} + A_1 \frac{\partial}{\partial x} \begin{bmatrix} u \\ v \\ \phi \end{bmatrix} + A_2 \frac{\partial}{\partial y} \begin{bmatrix} u \\ v \\ \phi \end{bmatrix} = 0,$$

with

$$A_1 = \begin{bmatrix} u_0 & 0 & 1 \\ 0 & u_0 & 0 \\ \phi & 0 & u_0 \end{bmatrix}, A_2 = \begin{bmatrix} v_0 & 0 & 0 \\ 0 & v_0 & 1 \\ 0 & \phi_0 & v_0 \end{bmatrix},$$

can be symmetrized with

$$H = \begin{bmatrix} \phi_0 & & \\ & \phi_0 & \\ & & 1 \end{bmatrix}.$$

The Variable Coefficient Case.

$$(2.1.25) \quad \frac{\partial u}{\partial t} = P(x, t, D)u.$$

This is the motivation for introducing the notion of hyperbolicity as is: freeze the coefficients and assume hyperbolicity of $u_t = P(x_0, t_0, D)u$ uniformly for each (x_0, t_0) ; then (unlike the case of weak hyperbolicity), the variable coefficients problem is also hyperbolic.

Remark: This result is based in the invariance of the notion of hyperbolicity under low-order perturbations; it restricts hyperbolic system to be of first-order.

So far we have dealt with hyperbolicity via the Fourier method studying the algebraic properties of its symbol $\hat{P}(ik)$; we can also work with the energy method.

For example, if we assume that $P(x, t, D)$ is semi-bounded, i.e., if

$$(2.1.26) \quad -M\|u\|_{L_2(x)}^2 \leq \operatorname{Re}(u, P(x, t, D)u)_{L_2(x)} \leq M\|u\|_{L_2(x)}^2, \quad 0 < M,$$

then we have hyperbolicity.

Example: The symmetric hyperbolic case $A_j(x, t) = A_j^T(x, t)$: we can rewrite such symmetric problems in the equivalent form

$$\frac{\partial u}{\partial t} = \frac{1}{2} \left[\sum_j A_j \frac{\partial u}{\partial x_j} + \sum_j \frac{\partial}{\partial x_j} (A_j u) \right] + Bu, \quad B = -\frac{1}{2} \sum_j \frac{\partial A_j}{\partial x_j}.$$

In this case the symmetry of the A_j 's implies that $\frac{1}{2} \left[\sum_j A_j \frac{\partial u}{\partial x_j} + \sum_j \frac{\partial}{\partial x_j} (A_j u) \right]$ is skew-adjoint, i.e., integration by parts gives

$$\left(u, \frac{1}{2} \left[\sum_j A_j \frac{\partial u}{\partial x_j} + \sum_j \frac{\partial}{\partial x_j} (A_j u) \right] \right)_{L_2(x)} \equiv 0.$$

Therefore we have

$$\operatorname{Re}(u, P(x, t, D)u)_{L_2(x)} \equiv \operatorname{Re}(Bu, u)_{L_2(x)},$$

and hence the semi-boundedness requirement (2.1.26) holds with $M = \|\operatorname{Re} B\|$. Consequently, if $A_j(x, t)$ are symmetric (or at least symmetrizable) then the system (2.1.16a) is hyperbolic.

2.2. Initial Value Problems of Parabolic Type.

The heat equation,

$$(2.2.1) \quad u_t = au_{xx}, \quad a > 0,$$

is the prototype of P.D.E.'s of parabolic type. We study the pure initial-value problem associated with (2.2.1), augmented with 2π -periodic boundary conditions and subject to initial conditions

$$(2.2.2) \quad u(x, 0) = f(x).$$

We can solve this equation using the Fourier method which gives

$$(2.2.3) \quad \hat{u}(k, t) = e^{-ak^2t} \hat{f}(k).$$

It shows the dissipation effect (= the rapid decay of the amplitudes, $|\hat{u}(k, t)|$, as functions of the high wavenumbers, $|k| \gg 1$) in this case, which is the essential feature of parabolicity.

As before, we study the manner in which the solution depends on its initial data.

1. Linearity: the principal of superposition holds.
2. Uniqueness: the solution is uniquely determined for $t > 0$ by the explicit formula

$$(2.2.4) \quad u(x, t) = \int_{y=-\infty}^{\infty} Q(x-y, t) f(y) dy, \quad Q(z) = \frac{1}{\sqrt{4\pi at}} e^{-\frac{z^2}{4at}} > 0.$$

3. Existence for large enough set of admissible initial data: bounded initial data $f(x)$ can be prescribed (or at least $|f(x)| \leq e^{\alpha x^2}$), and the corresponding solution is C^∞ - in fact $u(x, t > 0)$ is analytic because of exponential decay in Fourier space.
4. The maximum principle: follows directly from the representation of $u(x, t)$ as a convolution of $f(x)$ with the unit mass positive kernel $Q(z)$.
5. Energy decay: as usual we may proceed in one of two ways.

The Fourier Method. We start with

$$\left\| \frac{\partial^s}{\partial x^s} u(x, t) \right\|_{L_2}^2 \leq 2\pi \sum_k |\hat{f}(k)|^2 \cdot \text{Max}_k [|k|^{2s} \cdot |e^{-ak^2t}|^2] \leq \text{Const.} \cdot t^{-s} \cdot \|f\|_{L_2}^2,$$

(since the quantity inside the above brackets is maximized at $k^2 at = s$). The last a priori estimate shows that the parabolic solution becomes infinitely smoother than its initial data

is as we "gain" infinitely many s -derivatives, and at the same time, higher derivatives decay faster as $t \uparrow \infty$. Alternatively, we can work with the

Energy Method.

Multiply (2.2.1) by u and integrate to get

$$(2.2.6) \quad \frac{1}{2} \frac{d}{dt} \|u\|_{L_2(x)}^2 \leq -a \|u_x\|_{L_2(x)}^2$$

and in general

$$(2.2.7) \quad \frac{1}{2} \frac{d}{dt} \left\| \frac{\partial^s u}{\partial x^s} \right\|_{L_2}^2 \leq -\text{Const.} \left\| \frac{\partial^{s+1} u}{\partial x^{s+1}} \right\|_{L_2}^2$$

successive integration of (2.2.7) yields (2.2.5).

Turning to general case, we consider m th-order P.D.E.'s of the form,

$$(2.2.8) \quad \frac{\partial u}{\partial t} = P(x, t, D)u, \quad P(x, t, D) = \sum_{|j|=0}^m A_j(x, t) D^j.$$

We say that the system (2.2.8) is weakly parabolic of order α if

$$(2.2.9) \quad \left\| \frac{\partial^s}{\partial x^s} u(x, t) \right\|_{L_2} \leq \text{Const.} t^{-|s|/\alpha} \|u(x, 0)\|_{L_2(x)}.$$

In the constant coefficients case this leads to the Garding-Petrovski characterization of parabolicity of order β , requiring

$$\text{Re} \lambda \left[\hat{P}(ik) = \sum_{|j|=0}^m A_j(ik) D^j \right] \leq -A \cdot |k|^\beta + B.$$

Remark: Generically we have $\alpha = \beta = m$ the order of dissipation which is necessarily even.

The extension to the variable coefficients case (with Lipschitz continuous coefficients) may proceed in one of two ways. Either, we freeze the coefficients and apply the Fourier method to the constant coefficients problems; or, we may use the energy method, e.g., integration by parts shows that for

$$P(x, t, D) = \sum_j \frac{\partial}{\partial x_j} \left(A_j(x, t) \frac{\partial u}{\partial x_j} \right) + B_j \frac{\partial u}{\partial x_j} + Cu,$$

with $A_j + A_j^* > \delta > 0$, and $B_j = B_j^*$, the corresponding system (2.2.8) is parabolic of order 2.

Example: $u_t = au_{xx} + u_{xxx}$ is weakly parabolic of order two, yet does not satisfy Petrovski parabolicity.

2.3. Well-Posed Time-Dependent Problems.

Hyperbolic and parabolic equations are the two most important examples of time-dependent problems whose evolution process is well-posed. Thus, consider the initial value problem

$$(2.3.1) \quad \frac{\partial u}{\partial t} = P(x, t, D)u.$$

We assume that a large enough class of admissible initial data

$$(2.3.2) \quad u(x, t = 0) = f(x)$$

there exists a unique solution, $u(x, t)$. This defines a solution operator, $E(t, \tau)$ which describes the evolution of the problem

$$(2.3.3) \quad u(t) = E(t, \tau)u(\tau).$$

Hoping to compute such solutions, we need that the solutions will depend continuously in their initial data, i.e.,

$$(2.3.4) \quad \|u(t) - v(t)\| \leq \text{Const}_T \|u(0) - v(0)\|_{H^s} \quad 0 \leq t \leq T.$$

In view of linearity, this amounts to having the a priori estimate (boundedness)

$$(2.3.5) \quad \|u(t) \equiv E(t, \tau)u(\tau)\| \leq \text{Const}_T \|u(\tau)\|_{H^s}, \quad 0 \leq t \leq T,$$

which includes the hyperbolic and parabolic cases.

Counterexample: (Hadamard) By Cauchy-Kowaleski, the system

$$\frac{\partial u}{\partial t} + A \frac{\partial u}{\partial x} = 0, \quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix},$$

has a unique solution for arbitrary analytic data, at least for sufficiently small time. Yet, with initial data

$$(2.3.6) \quad u_1(x, 0) = \frac{\sin nx}{n}, \quad u_2(x, 0) = 0,$$

we obtain the solution

$$u_1(x, t) = \frac{\cosh nt \sin nx}{n}, \quad u_2(x, t) = \frac{\sinh nt \cos nx}{n}$$

which tends to infinity $\|u(\cdot, t)\|_{n \rightarrow \infty} \rightarrow \infty$, while the initial data tend to zero. Thus this Laplace equation, $\frac{\partial^2 u_1}{\partial t^2} + \frac{\partial^2 u_1}{\partial x^2} = 0$, is not well-posed as an initial-value problem.

Finally, we note that a well-posed problem is stable against perturbations of inhomogeneous

data in view of the following

Duhamel's principle. The solution of the inhomogeneous problem

$$(2.3.8) \quad \frac{\partial u}{\partial t} = P(x, t, D)u + F(x, t)$$

is given by

$$(2.3.9) \quad u(t) = E(t, 0)u(0) + \int_{\tau=0}^t E(t, \tau)F(\tau)d\tau.$$

Proof: We have

$$\begin{aligned} \frac{\partial}{\partial t}u(t) &= \frac{\partial}{\partial t}[E(t, 0)u(0)] + \frac{\partial}{\partial t} \left[\int_{\tau=0}^t E(t, \tau)F(\tau)d\tau \right] \\ &= P(x, t, D)[E(t, 0)u(0)] + E(t, t)F(t) + \int_{\tau=0}^t \frac{\partial}{\partial t}[E(t, \tau)F(\tau)]d\tau \\ &= P(x, t, D)[E(t, 0)u(0) + \int_{\tau=0}^t E(t, \tau)F(\tau)d\tau] + F(t) = P(x, t, D)u(t) + F. \end{aligned}$$

This implies the a priori estimate

$$(2.3.10) \quad \|u(t)\| \leq \text{Const}_T \|u(0)\| + \text{Const}_T \int_{\tau=0}^t \|F(\tau)\|_{H^s} d\tau, \quad 0 \leq t \leq T,$$

as asserted.

3. THE FOURIER METHOD FOR HYPERBOLIC AND PARABOLIC EQUATIONS

3.1. The Spectral Approximation

We begin with the simplest hyperbolic equation – the scalar constant-coefficients wave equation

$$(3.1.1) \quad \frac{\partial u}{\partial t} = a \frac{\partial u}{\partial x}$$

subject to initial conditions

$$(3.1.2) \quad u(x, 0) = f(x),$$

and periodic boundary conditions.

This Cauchy problem can be solved by the Fourier method: with $f(x) = \sum_{-\infty}^{\infty} \hat{f}(k) e^{ikx}$ we obtain after integration of (3.1.1),

$$(3.1.3) \quad \frac{\partial}{\partial t} \hat{u}(k, t) = ika \hat{u}(k, t),$$

with solution

$$(3.1.4) \quad \hat{u}(k, t) = e^{ikat} \hat{f}(k),$$

and hence

$$(3.1.5) \quad u(x, t) = \sum_k e^{ikat} \hat{f}(k) e^{ikx} = \sum_k \hat{f}(k) e^{ik(x+at)} = f(x + at).$$

Thus the solution operation in this case amounts to a simple translation

$$E(t, \tau)u(x, \tau) = u(x + a(t - \tau), t), \quad \|E(t, \tau)\| = 1.$$

This is reflected in the Fourier space, see (3.1.4), where each of the Fourier coefficients has the same change in phase and no change in amplitude; in particular, therefore, we have the a priori energy bound (conservation)

$$(3.1.6) \quad \|u(\cdot, t)\|^2 = 2\pi \sum_k |\hat{u}(k, t)|^2 = 2\pi \sum_k |\hat{f}(k)|^2 = \|f(\cdot)\|^2.$$

We want to solve this equation by the spectral Fourier method. To this end we shall approximate the spectral Fourier projection of the exact solution $Su_N \equiv S_N u(x, t)$. Projecting the equation (3.1.1) into the N -space we have

$$(3.1.7) \quad \frac{\partial u_N}{\partial t} = S_N \left[a \frac{\partial u}{\partial x} \right].$$

Since S_N commute with multiplication by a constant and with differentiation we can write this as

$$(3.1.8) \quad \frac{\partial u_N}{\partial t} = a \frac{\partial u_N}{\partial x}.$$

Thus $u_N = S_N u$ satisfies the same equation the exact solution does, subject to the approximate initial data

$$(3.1.9) \quad u_N(t = 0) = S_N f.$$

The resulting equations amount to $2N + 1$ ordinary differential equations (O.D.E.) for the amplitudes of the projected solution

$$(3.1.10) \quad \frac{d}{dt} \hat{u}_N(k, t) = ika \hat{u}_N(k, t), \quad -N \leq k \leq N,$$

subject to initial conditions

$$(3.1.11) \quad \hat{u}_N(k, 0) = \hat{f}(k).$$

Since these equations are independent of each other, we can solve them directly, obtaining

$$(3.1.12) \quad \hat{u}_N(k, t) = e^{ik a t} \hat{f}(k)$$

and our approximate solution takes the form

$$(3.1.13) \quad u_N(x, t) = \sum_{k=-N}^N \hat{f}(k) e^{ik(x+at)}.$$

Hence, the approximate solution $u_N(x, t) = f_N(x + at)$ satisfies

$$(3.1.14) \quad u(x, t) - u_N(x, t) = E(t, 0)f(x) - E(t, 0)S_N f(x)$$

and therefore, it converges spectrally to the exact solution, compares (1.1.26),

$$(3.1.15) \quad \begin{aligned} \|u(t) - u_N(t)\| &\leq \|E(t, 0)(I - S_N)f(x)\| \leq \\ &\leq \|(I - S_N)f(x)\| \leq \text{Const} \|f\|_{H^s} \cdot \frac{1}{N^s}. \end{aligned}$$

Similar estimates holds for higher Sobolev norms; in fact if the initial data is analytic then the convergence rate is exponential. In this case the only source of error comes from the initial data, that is we have the error equation

$$(3.1.16) \quad \frac{\partial}{\partial t}[u - u_N] = a \frac{\partial}{\partial x}[u - u_N]$$

subject to initial error

$$(3.1.17) \quad u - u_N(t=0) = f - f_N.$$

Consequently, the a priori estimate of this constant coefficient wave equation

$$(3.1.18) \quad \|u - u_N(t)\| \leq \text{Const}_T \|f - f_N\| \leq \text{Const} \cdot \|f\|_{H^s} \cdot \frac{1}{N^s} \quad \text{Const}_T = 1.$$

Now let us turn to the scalar equation with variable coefficients

$$(3.1.19) \quad \frac{\partial u}{\partial t} = a(x, t) \frac{\partial u}{\partial x}, \quad a(x, t) = 2\pi - \text{periodic}.$$

This hyperbolic equation is well-posed: by the energy method we have

$$(3.1.20) \quad \frac{1}{2} \frac{d}{dt} \int_x u^2(x, t) dx = \overbrace{-\int u_x a(x, t) u - \int a_x(x, t) u^2}^{\int u a(x, t) u_x} = -\frac{1}{2} \int a_x(x, t) u^2(x, t) dx,$$

and hence

$$(3.1.22a) \quad \|u(x, t)\|_{L_2(x)} \leq \text{Const}_T \cdot \|f(x)\|$$

with

$$(3.1.22b) \quad \text{Const}_T = e^{MT}, \quad M = \text{Max}_{x,t} [-a_x(x, t)].$$

In other words, we have for the solution operator

$$\|S(t, \tau)u(\tau)\|_{L_2(x)} \leq e^{M(t-\tau)} \|u(\tau)\|_{L_2(x)}$$

and similarly for higher norms. As before, we want to solve this equation by the spectral Fourier method. We consider the spectral Fourier projection of the exact solution $u_N = S_N u(x, t)$; projecting the equation (3.1.19) we get

$$(3.1.23) \quad \frac{\partial}{\partial t} u_N = S_N \left[a(x, t) \frac{\partial u}{\partial x} \right].$$

Unlike the previous constant coefficients case, now S_N does not commute with multiplication by $a(x, t)$, that is, for arbitrary smooth function $\rho(x, t)$ we have (suppressing time dependence)

$$(3.1.24a) \quad S_N a(x) \rho(x) = \sum_{k=-N}^N \left(\sum_{j=-\infty}^{\infty} \hat{a}(k-j) \hat{\rho}(j) \right) e^{ikx}$$

while

$$(3.1.24b) \quad a(x)S_N\rho(x) = \sum_{k=-\infty}^{\infty} \left(\sum_{j=-N}^N \hat{a}(k-j)\hat{\rho}(j) \right) e^{ikx}.$$

Thus, if we exchange the order of operations we arrive at

$$(3.1.25) \quad \frac{\partial u_N}{\partial t} = a(x, t) \frac{\partial u_N}{\partial x} - [a(x, t)S_N - S_N a(x, t)] \frac{\partial u}{\partial x}.$$

While the second term on the right is not zero, this commutator between multiplication and Fourier projection is spectrally small, i.e.,

$$(3.1.26) \quad \begin{aligned} & \|S_N a(x)\rho(x) - a(x)S_N\rho(x)\|_{L_2(x)} = \\ & \| (S_N - I)a(x)\rho(x) + a(x)(I - S_N)\rho(x) \|_{L_2(x)} \leq \\ & \leq \text{Const.} \|a(x)\rho(x)\|_{H^s} \cdot \frac{1}{N^s} + \text{Const.} \|a(x)\|_{L_\infty(x)} \cdot \|\rho(x)\|_{H^s} \cdot \frac{1}{N^s} \end{aligned}$$

and so we intend to neglect this spectrally small contribution and to set as an approximate model equation for the Fourier projection of $u(x, t)$

$$\frac{\partial v_N}{\partial t} = a(x, t) \frac{\partial v_N}{\partial x}.$$

Yet, the second term may lie outside the N-space, and so we need to project it back, thus arriving at our final form for the spectral Fourier approximation of (3.1.19)

$$(3.1.27) \quad \frac{\partial v_N}{\partial t} = S_N \left(a(x, t) \frac{\partial v_N}{\partial x} \right).$$

Again, we commit here a spectrally small deviation from the previous model, for

$$(3.1.28) \quad \|(I - S_N)a\rho(x)\|_{L_2(x)} \leq \text{Const} \|a(x)\rho(x)\|_{H^s} \cdot \frac{1}{N^s}.$$

The Fourier projection of the exact solution does not satisfy (3.1.22), but rather a near by equation,

$$(3.1.29) \quad \frac{\partial u_N}{\partial t} = S_N \left(a(x, t) \frac{\partial u_N}{\partial x} \right) + F_N(x, t)$$

where the truncation error, $F_N(x, t)$ is given by

$$(3.1.30) \quad F_N(x, t) = S_N \left[a(x, t)(I - S_N) \frac{\partial u}{\partial x} \right].$$

The truncation error is the amount by which the (projection of) the exact solution misses our approximate mode (3.1.27); in this case it is spectrally small by the errors committed in (3.1.26) and (3.1.18). More precisely we have

$$(3.1.31) \quad \|F_N(x, t)\|_{L_2(x)} \leq \|a(x, t)\|_{L_2(x)} \cdot \|u\|_{H^{s+1}} \frac{1}{N^s},$$

depending on the degree of smoothness of the exact solution. We note that by hyperbolicity, the later is exactly the degree of smoothness of the initial data, i.e., by the hyperbolic differential energy estimate

$$(3.1.32) \quad \|F_N(x, t)\|_{L_2(x)} \leq \|a(x, t)\|_{L_2(x)} \cdot \|f\|_{H^{s+1}} \cdot \frac{1}{N^s}$$

and in the particular case of analytic initial data, the truncation error is exponentially small.

From this point of view, our spectral approximation (3.1.27) satisfies an evolution model which is spectrally away from that of the Fourier projection of the exact solution (3.1.29). This is in addition to the spectrally small error we commit initially, as we had before

$$(3.1.33) \quad v_N(t=0) = S_N f \equiv f_N.$$

We now raise the question of convergence. That is, whether the accumulation of spectrally small errors while integrating (3.1.27) rather than (3.1.29), give rise to an approximate solution $v_N(x, t)$ which is only spectrally away from the exact projection $u_N(x, t)$. We already know that the distance between $u_N(x, t)$ and the exact solution $u(x, t)$ – due to the spectrally small initial error – is spectrally small as we have seen in the previous constant coefficient case.

To answer this convergence question we have to require the stability of the approximate model (3.1.27). That is, we say that the approximation (3.1.27) is stable if it satisfies an a priori energy estimate analogous to the one we have for the differential equation

$$(3.1.34) \quad \|v_N(t)\| \leq \text{Const.} e^{Mt} \|v_N(0)\|.$$

Clearly, such a stability estimate is necessary in any computational model. Otherwise, the evolution model does not depend continuously on the (initial) data, and small rounding errors can render the computed solution useless. And on the positive side we will show that the stability implies the spectral convergence of an approximate solution $u_N(x, t)$.³ Indeed the error equation for $e_N(t) = u_N(t) - v_N(t)$ takes the form

$$(3.1.35) \quad \frac{\partial e_N}{\partial t} = S_N \left[a(x, t) \frac{\partial e_N}{\partial x} + F_N(x, t) \right].$$

³We note that in the previous constant coefficient case, the approximate model coincides with the differential case, hence the stability estimate was nothing but the a priori estimate for the differential equation itself.

Let $E_N(t, \tau)$ denote the evolution operator solution associated with our approximate model. By the stability estimate (3.1.34)

$$(3.1.36) \quad \|E_N(t, \tau)v_N(\tau)\| \leq \text{Const}e^{M(t-\tau)}\|v_N(\tau)\|.$$

Hence, by (3.1.36) together with Duhammel's principle we get for the inhomogeneous error equation (3.1.35)

$$(3.1.37a) \quad e_N(t) = E_N(t, 0)e_N(0) + \int_{\tau=0}^t E_N(t, \tau)F_N(\tau)d\tau$$

and

$$(3.1.37b) \quad \|e_N(t)\| \leq \text{Const}.e^{Mt} \left[\|e_N(0)\|_{L_2(x)} + \int_{\tau=0}^t \|F_N(x, \tau)\|_{L_2(x)}d\tau \right].$$

In our case $e_N(0) = f_N - Sf_N = 0$, and the truncation error $F_N(x, \tau)$ is spectrally small; hence

$$(3.1.38) \quad \|e_N \equiv u_N(t) - v_N(t)\| \leq \text{Const}.e^{Mt} \cdot \frac{1}{N^s}$$

where the constant depends on $\|a(x, t)\|_{L_\infty(!)}$ and $\|f\|_{H^{s+1}}$, i.e., restricted solely by the smoothness of the data. In the particular case of analytic data we have exponential convergence

$$(3.1.39) \quad \|e_N(t) \equiv u_N(t) - v_N(t)\| \leq \text{Const}.e^{Mt} \cdot e^{-\eta N}.$$

Adding to this the error between $u_N(t)$ and $u(t)$ (– which is due to the spectrally small error in the initial data between f_N and f) we end up with

$$(3.1.40) \quad \|u(t) - v_N(t)\| \leq \text{Const}.e^{Mt} \cdot \begin{cases} \frac{1}{N^s} & \text{for } H^{s+1} \text{ initial data} \\ e^{-\eta N} & \text{for analytic initial data} \end{cases}$$

To summarize, we have shown that our spectral Fourier approximation converges spectrally to the exact solution, provided the approximation (3.1.27) is stable.

Is the approximation (3.1.27) stable? That is, do we have the a priori estimate (3.1.34)? To show this we try to follow the steps that lead to the analogue estimate in the differential case, compare (3.1.20). Thus, we multiply (3.1.27) by $v_N(x, t)$ and integrate over the in-period, obtaining

$$(3.1.41) \quad \frac{1}{2} \frac{d}{dt} \int_x v_N^2(x, t)dx = + \int_x v_N(x, t)S_N \left(a(x, t) \frac{\partial v_N}{\partial x} \right) dx.$$

But $v_N(x, t)$ is orthogonal to $(I - S_N) \left[a(x, t) \frac{\partial v_N}{\partial x} \right]$ so adding this to the right-hand side of (3.1.41) we arrive at

$$(3.1.42) \quad \frac{1}{2} \frac{d}{dt} \int_x v_N^2(x, t) = \int_x v_N(x, t) a(x, t) \frac{\partial v_N}{\partial x} dx$$

and we continue precisely as before to conclude, similarly to (3.1.22), that the stability estimate (3.1.34) holds

$$(3.1.43) \quad \|v_N(t)\| \leq \text{Const.} e^{Mt} \|v_N(0)\|, \quad M = \text{Max}_{x,t} [-a_x(x, t)].$$

In the constant coefficient case the Fourier method amount to a system of $(2N + 1)$ uncoupled O.D.E.'s for the Fourier coefficients of $v_N = u_N$ which were integrated explicitly. Let's see what is the case with problems having variable coefficients say, for simplicity, $a \equiv a(x)$. Fourier transform (3.1.22) we obtain for $\hat{v}(k, t) = \hat{v}_N(k, t)$ - the k th-Fourier coefficient of $v_N(x, t) = \sum_{k=-N}^N \hat{v}(k, t) e^{ikx}$,

$$(3.1.44) \quad \frac{d\hat{v}(k, t)}{dt} = \sum_{j=-N}^N \hat{a}(k-j) ij \hat{v}(j, t), \quad -N \leq k \leq N.$$

In this case we have a $(2N + 1) \times (2N + 1)$ coupled system of O.D.E.'s written in the matrix-vector form, consult (1.2.46)

$$(3.1.45) \quad \frac{d}{dt} \hat{v}(t) = \hat{A} \Lambda \hat{v}(t), \quad \hat{v}(t) = \begin{bmatrix} \hat{v}(-N, t) \\ \vdots \\ \hat{v}(N, t) \end{bmatrix} \quad \hat{A}_{kj} = \hat{a}(k-j), \Lambda = \text{diag}(ik).$$

We can solve this system explicitly (since a (\cdot) was assumed not to depend on time)

$$(3.1.46) \quad \hat{v}(t) = e^{\hat{A} \Lambda t} \hat{v}(0);$$

that is, we obtain an explicit representation of the solution operator

$$(3.1.47) \quad E_N(t, \tau) = F_N^{-1} e^{\hat{A} \Lambda (t-\tau)} F_N, \quad \hat{A} = \hat{A}_N, \Lambda = \Lambda_N$$

where F_N denote the spectral Fourier projection

$$(3.1.48) \quad F_N v_N(x) = \begin{bmatrix} \hat{v}(-N) \\ \vdots \\ \hat{v}(N) \end{bmatrix}.$$

We note that in view of Parseval identity $\|F_N v_N(x)\|_2 = \|v_N(x)\|_{L_2(x)}$ (modulo factorization factor), hence, stability amounts to having the a priori estimate on the discrete symbol $\hat{E}_N(t, \tau) = e^{\hat{A}_N \Lambda (t-\tau)}$, requiring

$$(3.1.49) \quad \|e^{\hat{A}_N \Lambda (t-\tau)}\| \leq \text{Const.} e^{M(t-\tau)}.$$

The essential point of stability here, lies in having a uniform bound in the RHS of (3.1.49) independent on the order of the system, e.g., a straightforward estimate of the form

$$(3.1.50) \quad \|e^{\hat{A}_N \Lambda(t-\tau)}\| \leq e^{\|\hat{A}_N\| \cdot \|\Lambda\|(t-\tau)}$$

will not do because $\|\Lambda_N\|_{N \rightarrow \infty} \uparrow \infty$. The essence of the a priori estimate we obtained in (3.1.22), and likewise in (3.1.42), was that the (unbounded) operator $P(x, t, D) \equiv a(x, t) \partial_x$ is semi-bounded, i.e.,

$$\operatorname{Re} \left[a(x, t) \frac{\partial}{\partial x} \right] = \frac{1}{2} \left[a(x, t) \frac{\partial}{\partial x} - \frac{\partial}{\partial x} (a(x, t) \cdot) \right] = -\frac{1}{2} a_x(x, t);$$

namely, compare (2.1.26)

$$\left(\operatorname{Re} \left[a(x, t) \frac{\partial}{\partial x} \right] u, u \right)_{L_2(x)} \leq M \|u\|_{L_2(x)}^2$$

and likewise for $\operatorname{Re} \left(S_N \left[a(x, t) \frac{\partial}{\partial x} \right] \right)$. In the present form this is expressed by the sharper estimate of the matrix exponent,⁴ compare (3.1.50)

$$(3.1.51) \quad \|e^{\hat{A}_N \Lambda(t-\tau)}\| \leq e^{\|\operatorname{Re} \hat{A}_N \Lambda\| \cdot (t-\tau)}.$$

This time, $\|\operatorname{Re} \hat{A}_N \Lambda\|$ like the $\operatorname{Re}[P(x, t, D)]$, is bounded. Indeed, $[\operatorname{Re} \hat{A} \Lambda]_{kj} = \frac{1}{2} [\hat{a}(k-j)ij + \overline{\hat{a}(j-k)ik}]$, and since $a(x, t)$ is real (hyperbolicity!) then $\overline{\hat{a}(p)} = \hat{a}(-p)$, i.e.,

$$(3.1.52) \quad [\operatorname{Re} \hat{A} \Lambda]_{kj} = \frac{1}{2} i(j-k) \hat{a}(k-j) \quad -N \leq j, k \leq N.$$

Thus, $\operatorname{Re} \hat{A} \Lambda$ is a Toeplitz matrix, namely its (k, j) entry depends solely on its distance from the main diagonal $k-j$; we leave it as an exercise – using our previous study on circulant matrices in (1.2.43) – to see that its norm does not exceed the sum of absolute values along the, say, zeroth ($j=0$) row, i.e.,

$$(3.1.53) \quad \|\operatorname{Re} \hat{A}_N \Lambda\| \leq \frac{1}{2} \sum_{k=-N}^N |k \hat{a}(k)|$$

which is bounded, uniformly with respect to N , provided $a(x, t)$ is sufficiently smooth, e.g., we can take the exponent M to be

$$(3.1.54) \quad \begin{aligned} M = \frac{1}{2} \sum_{k=-N}^N |k \hat{a}(k)| &\leq \frac{1}{2} \sqrt{\sum k^4 |a(k)|^2} \cdot \sum_{k=-N}^N \frac{1}{k^2} \leq \\ &\leq \frac{\pi}{6} \cdot \|a_{xx}(x, t)\|_{L^2(x)} \end{aligned}$$

⁴To see this, use Duhammel's principle for $\frac{d\hat{v}}{dt} = \operatorname{Re} \hat{A} \Lambda \hat{v}(t) + F(t)$ where $F(t) = i \operatorname{Im} \hat{A} \Lambda e^{\hat{A} \Lambda t}$ or integrate directly.

which is only slightly worse than what we obtained in (3.1.43).

A similar analysis shows the convergence of the spectral-Fourier method for hyperbolic systems. For example, consider the $N \times N$ symmetric hyperbolic problem

$$(3.1.55a) \quad \frac{\partial u}{\partial t} = A(x, t) \frac{\partial u}{\partial x} + B(x, t)u, \quad \text{with symmetric } A(x, t).$$

We note that if the system is not in this symmetric form, then (in the 1-D case) we can bring it to the symmetric form by a change of variables, i.e., the existence of a smooth symmetric $H(x, t)$ such that $H(x, t)A(x, t)$ is symmetric, implies that for $w(x, t) = T^{-1}(x, t)u(x, t)$ with $H = (T^{-1})^*T^{-1}$ we have, compare (2.1.15b)

$$(3.1.55b) \quad \frac{\partial w}{\partial t} = T^{-1}(x, t)A(x, t)T(x, t) \frac{\partial w}{\partial x} + C(x, t)w(x, t)$$

where $T^{-1}(x, t)A(x, t)T(x, t) \equiv T^*(x, t)H(x, t)A(x, t)T(x, t)$ is symmetric, and $C(x, t) = B(x, t) + \frac{\partial T^{-1}}{\partial t}(x, t) - T^{-1}(x, t)A(x, t)\frac{\partial T}{\partial x}(x, t)$. The spectral Fourier approximation of (3.1.55a) takes the form

$$(3.1.56) \quad \frac{\partial v_N}{\partial t} = S_N \left(A(x, t) \frac{\partial u_N}{\partial x} \right) + S_N B(x, t)v_N(x, t).$$

Its stability follows from integration by parts, for by orthogonality

$$(3.1.57a) \quad \frac{1}{2} \frac{d}{dt} \int_x v_N^2(x, t) dx = \int v_N A(x, t) \frac{\partial v_N}{\partial x} dx + \int u_N B(x, t) u_N dx \leq M \int_x v_N^2(x, t) dx$$

where

$$(3.1.57b) \quad M = \text{Max}_{x,t} \left[-\frac{\partial A(x, t)}{\partial x} + \text{Re} B(x, t) \right]$$

and hence

$$(3.1.58) \quad \|v_N(t)\|_{L^2(x)} \leq e^{Mt} \|v_N(0)\|.$$

The approximation (3.1.56) is spectrally accurate with (3.1.55) and hence spectral convergence follows. The solution of (3.1.56) is carried out in the Fourier space, and takes the form

$$(3.1.59) \quad \frac{d}{dt} \hat{v}(k, t) = \sum_{j=-N}^N \hat{A}(k-j, t) i j \hat{v}(j, t), \quad -N \leq k \leq N,$$

which form a coupled $(2N+1) \times (2N+1)$ system of O.D.E.'s for the $(2N+1)$ -vectors of Fourier coefficients $\hat{v}(k, t)$.

There are two difficulties in carrying out the calculation with the spectral Fourier method. First, is the time integration of (3.1.59); even in the constant coefficient case, it requires to

compute the exponent $e^{\hat{A}\Lambda t}$ which is expensive, and in the time-dependent case we must appeal to approximate numerical methods for time integration. Second, to compute the RHS of (3.1.59) we need to multiply an $(2N + 1) \times (2N + 1)$ matrix, $\hat{A}\Lambda$ by the Fourier coefficient vector which requires $\mathcal{O}(N^2)$ operations. Indeed, since \hat{A} is a Toeplitz matrix and Λ is diagonal, we can still carry out this multiplication efficiently, i.e., using two FFT's which requires $\mathcal{O}(N \log N)$ operations. Yet, it still necessitates to carry out the calculation in the Fourier space. We can overcome the last difficulty with the pseudospectral Fourier method.

Before leaving the spectral method, we note that its spectral convergence equally applies to any P.D.E.

$$(3.1.60) \quad \frac{\partial u}{\partial t} = P(x, t, D)u$$

with semi-bounded operator $P(x, t, D)$, e.g., the symmetric hyperbolic as well as the parabolic operators. Indeed, the spectral approximation of (3.1.60) reads

$$(3.1.61) \quad \frac{\partial v_N}{\partial t} = S_N P(x, t, D)v_N.$$

Multiply by v_N and integrate – by orthogonality and semi-boundedness we have

$$(3.1.62) \quad \frac{1}{2} \frac{d}{dt} \int_{\mathbf{x}} v_N^2(x, t) dx = \text{Re}(v_N, P(x, t, D)v_N) \leq M \int_{\mathbf{M}} v_N^2(x, t) dx.$$

Hence stability follows and the method converges spectrally.

3.2. The Pseudospectral Approximation

We return to the scalar constant coefficient case

$$(3.2.1) \quad \frac{\partial u}{\partial t} = a \frac{\partial u}{\partial x}$$

subject to periodic boundary conditions and prescribed initial data

$$(3.2.2) \quad u(x, 0) = f(x).$$

To solve this problem by the pseudospectral Fourier method, we proceed as before, this time projecting (3.2.1) with the pseudospectral projection ψ_N , to obtain for $u_N = \psi_N u(x, t)$

$$(3.2.3) \quad \frac{\partial u_N}{\partial t} = \psi_N \left(a \frac{\partial u}{\partial x} \right).$$

Here, ψ_N commutes with multiplication by a constant, but unlike the spectral case, it does not commute with differentiation, i.e., by the aliasing relation (1.2.2) we have

$$(3.2.4a) \quad \psi_N \frac{\partial \rho}{\partial x} \sum_{k=-N}^N (k \tilde{\rho}(k)) e^{ikx} = \sum_{k=-N}^N \sum_j i[k + j(2N+1)] \hat{\rho}[k + j(2N+1)] e^{ikx}$$

where as

$$(3.2.4b) \quad \frac{\partial}{\partial x} \psi_N \rho = \sum_{k=-N}^N (k \tilde{\rho}(k)) e^{ikx} = \sum_{k=-N}^N ik \sum_j \hat{\rho}[k + j(2N+1)] e^{ikx}.$$

The difference between these two operators is a pure aliasing error, i.e., we have for $\psi_N = S_N + A_N$, see (1.2.13)

$$\psi_N \frac{d\rho}{dx} - \frac{d}{dx}(\psi_N \rho) \equiv \left[A_N, \frac{d}{dx} \right] \rho = \sum_{k=-N}^N \sum_{j \neq 0} i[k + j(2N+1)] \hat{\rho}(k + j(2N+1)) e^{ikx}$$

which is spectrally small. Sacrificing such spectrally small errors, we are led to the pseudospectral approximation of (3.2.1)

$$(3.2.5) \quad \frac{\partial v_N}{\partial t} = a \frac{\partial v_N}{\partial x}$$

subject to initial conditions

$$(3.2.6) \quad v_N(t=0) = \psi_N f.$$

Here, $v_N = v_N(x, t)$ is an N-degree trigonometric polynomial which satisfies a nearby equation satisfied by the interpolant of the exact solution $\psi_N u(x, t)$. That is, $u_N \equiv \psi_N u(x, t)$ satisfies (3.2.5) modulo spectrally small truncation error

$$(3.2.7) \quad \frac{\partial u_N}{\partial t} = a \frac{\partial u_N}{\partial x} + F_N(x, t), \quad F_N(x, t) = a \psi_N \left[\frac{\partial}{\partial x} (I - \psi_N) u \right]$$

where by (3.2.3), $F_N(x, t) = a \left[\psi_N \frac{\partial u}{\partial x} - \frac{\partial}{\partial x}(\psi_N u) \right]$, and by (1.2.17) it is indeed spectrally small

$$(3.2.8) \quad \|F_N(x, t)\| \leq |a| \left\| \frac{\partial}{\partial x}[(I - \psi_N)u] \right\| \leq |a| \|u\|_{H^{s+1}} \frac{1}{N^s}.$$

The stability proof of (3.2.5) follows along the lines of the spectral stability, and spectral convergence follows using Duhammel's principle for the stable numerical solution operator. That is, the error equation for $e_N = u_N - v_N$ is

$$(3.2.9) \quad \frac{\partial e_N}{\partial t} = a \frac{\partial e_N}{\partial x} + F_N(x, t)$$

whose solution is

$$(3.2.10) \quad e_N(t) = E_N(t, 0)(f_N - \psi_N f) + \int_{\tau=0}^t E_N(t, \tau) F_N(x, \tau) d\tau.$$

Hence, by stability

$$(3.2.11) \quad \|e_N(t)\| \leq \text{Const.} e^{Mt} \cdot \|u\|_{H^{s+1}} \frac{1}{N^s} \leq \text{Const.} e^{Mt} \|f\|_{H^{s+1}} \cdot \frac{1}{N^s};$$

this together with the estimate of the pseudospectral projection yields

$$(3.2.12) \quad \|u(t) - v_N(t)\| \leq \text{Const.} e^{Mt} \cdot \begin{cases} \frac{1}{N^s} & \text{for } H^{s+1} \text{ initial data} \\ e^{-\eta N} & \text{for analytic initial data} \end{cases}$$

To carry out the calculation of (3.2.5) we can compute the discrete Fourier coefficients $\tilde{v}(k, t)$ which obey the O.D.E.,

$$(3.2.13) \quad \frac{d\tilde{v}}{dt}(k, v) = ika\tilde{v}(k, t),$$

as was done with the spectral case; alternatively, we can realize our approximate interpolant $v_N(x, t)$ at the $2N+1$ equidistant points $x_\nu = \nu h$, and (3.2.5) amounts to a coupled $(2N+1)$ - O.D.E. system in the real space

$$(3.2.14) \quad \frac{dv_N}{dt}(x_\nu, t) = a \frac{\partial v_N}{\partial x}(x = x_\nu, t) \quad \nu = 0, 1, \dots, 2N.$$

$$(3.2.15) \quad v_N(x_\nu, 0) = f(x_\nu).$$

Let us turn to the variable coefficient case,

$$(3.2.16) \quad \frac{\partial u}{\partial t} = a(x, t) \frac{\partial u}{\partial x}.$$

The pseudospectral approximation takes the form

$$(3.2.17a) \quad \frac{\partial v_N}{\partial t} = \psi_N \left[a(x, t) \frac{\partial v_N}{\partial x} \right]$$

$$(3.2.17b) \quad v_N(x_\nu, 0) = f(x_\nu).$$

It can be solved as a coupled O.D.E. system in the Fourier space, but just as easily can be realized at the $2N + 1$ so-called collocation points

$$(3.2.18a) \quad \frac{dv_N(x_\nu, t)}{dt} = a(x_\nu, t) \frac{\partial v_N}{\partial x}(x = x_\nu, t)$$

$$(3.2.18b) \quad v_N(x_\nu, t = 0) = f(x_\nu).$$

The truncation error of this model is spectrally small in the sense that $u_N = \psi_N u$ satisfies

$$(3.2.19) \quad \frac{\partial u_N}{\partial t} = \psi_N \left[a(x, t) \frac{\partial u_N}{\partial x} \right] + F_N(x, t)$$

where

$$(3.2.20) \quad F_N(x, t) = \psi_N \left[a(x, t) \frac{\partial u}{\partial x} \right] - \psi_N \left[a(x, t) \frac{\partial}{\partial x} (\psi_N u) \right]$$

is spectrally small

$$(3.2.21) \quad \begin{aligned} \|F_N(x, t)\| &\leq \|\psi_N \left[a(x, t) \frac{\partial}{\partial x} [(I - \psi_N)u] \right]\| \leq \\ &\leq \|a(x, t)\|_{L^\infty} \cdot \|f\|_{H^{s+1}} \cdot \frac{1}{N^s}. \end{aligned}$$

Hence, if the approximation (3.2.12) stable, spectral convergence follows. Is the approximation (3.2.12) stable? The presence of aliasing errors is responsible to a considerable difficulty in proving this, and currently this is an open question. If we try to follow the differential and spectral recipe, we should multiply by $v_N(x, t)$ and integrate by parts. However, here $v_N(x, t)$ is not orthogonal to $(I - \psi_N)[\dots]$ which otherwise would enable us to follow the differential estimate of $\int v_N(x, t) a(x, t) \frac{\partial v_N}{\partial x}(x, t) dx$ in terms of $\int_x v_N^2(x, t) dx$; more precisely, we have for $\psi_N = S_N + A_N$ that $I - \psi_N = I - S_N - A_N$ where $\int v_N(I - S_N)[\dots] dx = 0$, compare (3.1.41), (3.1.42); yet $\int v_N A_N[\dots] dx$ leaves us with an additional contribution which cannot be bounded in terms of $\int_x v_N^2(x, t) dx$ in order to end up with the stability proof with Gronwall's inequality. To shed a different light on this difficulty, we can turn to the Fourier space; we write (3.2.17) in the form

$$(3.2.22) \quad \frac{\partial v_N}{\partial t} = a(x) \frac{\partial v_N}{\partial x}$$

and Fourier transform to get for the k th Fourier coefficient

$$(3.2.23a) \quad \frac{d}{dt} \tilde{v}(k, t) = \sum_{j=-N}^N \tilde{a}(k-j, t) i j \tilde{v}(j, t)$$

i.e.,

$$(3.2.23b) \quad \frac{d}{dt} \tilde{v}(t) = \tilde{A}_N \Lambda \tilde{v}(t) \quad \tilde{A}_{kj} = \sum_p \hat{a}[k-j+p(2N+1)].$$

This time, $\text{Re} \tilde{A}_N \Lambda$ is unbounded. This difficulty appears when we confine ourselves to the discrete framework: multiplying (3.2.18a) by $v(x_\nu, t)$ and trying to sum by parts we arrive at

$$(3.2.24) \quad \begin{aligned} \frac{1}{2} \frac{d}{dt} \sum_\nu v_N^2(x_\nu, t) &= \sum_\nu a(x_\nu, t) v(x_\nu, t) \frac{\partial v}{\partial x}(x_\nu, t) \\ &= \sum_\nu \frac{\partial}{\partial x} \left[\frac{1}{2} a(x, t) v^2(x, t) \right] \Big|_{x=x_\nu} - \sum_\nu \frac{1}{2} a'(x_\nu, t) v_N^2(x_\nu, t); \end{aligned}$$

but the first term on the right does not vanish in this case – it equals, by the aliasing relation, to

$$\sum_\nu \frac{\partial}{\partial x} \left[\frac{1}{2} a(x, t) v^2(x, t) \right] \Big|_{x=x_\nu} = \int \frac{\partial}{\partial x} [\cdots] + \sum_{p \neq 0} p \cdot (2N+1) \frac{1}{2} \hat{a} v^2[p \cdot (2N+1)]$$

and a loss of one derivative is reflected by the factor $2N+1$ inside the right summation. There are two main approaches to enforce stability at this point: skew-symmetric differencing and smoothing. We discuss these approaches in the next two subsections.

3.3 Skew-Symmetric Differencing The essential argument of well-posedness for symmetric hyperbolic systems with constant coefficients is the fact that (say in the 1-D case) $P(D) = A \frac{\partial}{\partial x}$ is a skew-adjoint operator what is loosely called “integration by parts”..... With variable coefficients this is also true, modulo low-order bounded terms, i.e.,

$$(3.3.1) \quad P(x, t, D) \equiv A(x, t) \frac{\partial}{\partial x} = \frac{1}{2} \left[A(x, t) \frac{\partial}{\partial x} + \frac{\partial}{\partial x} (a(x, t) \cdot) \right] - \frac{1}{2} A_x(x, t).$$

The stability proofs of spectral methods follow the same line, i.e., we have in the Fourier space, compare (3.1.45), (3.1.52)

$$(3.3.2) \quad \hat{A}_N \Lambda = \frac{1}{2} [\hat{A}_N \Lambda - \Lambda \hat{A}_N] + \frac{1}{2} [\hat{A}_N \Lambda + \Lambda^* \hat{A}_N]$$

and stability amounts to show that the second term in (3.3.2) is bounded: for then we have in (3.3.2) (as in (3.3.1)) a skew-adjoint term with an additional bounded operator. The difficulty with the stability of pseudo-spectral methods arises from the fact that the second term on the right of (3.3.2) is unbounded,

$$(3.3.3) \quad \lim_{N \rightarrow \infty} \left\| \frac{1}{2} (\tilde{A}_N \Lambda + \Lambda^* \tilde{A}_N) \right\| \uparrow \infty.$$

To overcome this difficulty, we can discretize the symmetric hyperbolic system (again, say the 1-D case)

$$(3.3.4a) \quad \frac{\partial u}{\partial t} = A(x, t) \frac{\partial u}{\partial x}$$

when the spatial operator is already put in the “right” skew-adjoint form, compare (3.3.1),

$$(3.3.4b) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \left[A(x, t) \frac{\partial u}{\partial x} + \frac{\partial}{\partial x} (A(x, t) u) \right] - \frac{1}{2} A_x(x, t) u.$$

The pseudospectral approximation takes the form

$$(3.3.5) \quad \frac{\partial u_N}{\partial t} = \frac{1}{2} \left[\psi_N \left(A(x, t) \frac{\partial u_N}{\partial x} \right) + \frac{\partial}{\partial x} \psi_N (A(x, t) u_N) \right] - \frac{1}{2} \psi_N (A_x(x, t) u_N).$$

In the Fourier space, this gives us

$$(3.3.6) \quad \frac{d\tilde{v}}{dt} = \frac{1}{2} [\tilde{A}_N \Lambda + \Lambda \tilde{A}_N] \tilde{v} - \frac{1}{2} \frac{\partial A_N}{\partial x} \tilde{v}.$$

Now, $\tilde{A}_N \Lambda + \Lambda \tilde{A}_N$ is symmetric because Λ is, $\frac{\partial A_N}{\partial x}$ is bounded and stability follows.

3.4. Smoothing

We have already met the process of smoothing in connection with the heat equation: starting with bounded initial data, $f(x)$, the solution of the heat equation (2.2.1)

$$(3.4.1) \quad u(x, t) = Q * f(x), \quad Q(x) = \frac{1}{\sqrt{4\pi a}} e^{-\frac{x^2}{4at}}, \quad t > 0$$

represents the effect of smoothing $f(x)$, so that $u(\cdot, t > 0) \in C^\infty$ (in fact analytic) and $u(x, t \downarrow 0) = f(x)$.

A general process of smoothing can be accomplished by convolution with appropriate smoothing kernel $Q_\varepsilon(x)$

$$(3.4.2) \quad f_\varepsilon(x) = Q_\varepsilon(x) * f(x)$$

such that

$$(3.4.3a) \quad Q_\varepsilon(x) * f(x)$$

is sufficiently smoother than $f(x)$ is, and

$$(3.4.3b) \quad Q_\varepsilon(x) * f(x) \xrightarrow{\varepsilon \rightarrow 0} f(x).$$

With the heat kernel, the role of ε was played by time $t > 0$. A standard way to construct such smoothers is the following. We start with a C^∞ -function supported on, say, $(-1, 1)$, such that it has a unit mass

$$(3.4.4a) \quad \int_{-1}^1 Q(x) dx = 1$$

and zero first r moments

$$(3.4.4b) \quad \int_{-1}^1 x^j Q(x) dx = 0, \quad j = 1, 2, \dots, r.$$

Then we set $Q_\varepsilon(x) = \frac{1}{\varepsilon} Q\left(\frac{x}{\varepsilon}\right)$ and consider

$$(3.4.5) \quad f_\varepsilon(x) = Q_\varepsilon(x) * f(x), \quad \varepsilon > 0.$$

Now, assume f is $(r + 1)$ -differentiable in the ε neighborhood of x ; then, since $Q_\varepsilon(x)$ is supported on $(-\varepsilon, \varepsilon)$ and satisfies (3.4.4a) as well, we have by Taylor expansion

$$(3.4.6) \quad \begin{aligned} f(x) - Q_\varepsilon(x) * f(x) &= \int_{|y| \leq \varepsilon} Q_\varepsilon(y) [f(x) - f(x - y)] dy = \\ &= \int_{|y| \leq \varepsilon} Q_\varepsilon(y) \left[\sum_{j=1}^r \frac{(-y)^j}{j!} f^{(j)}(x) + \frac{(-y)^{r+1}}{(r+1)!} f^{(r+1)}(\xi) \right] dy. \end{aligned}$$

The first r moments of $Q_\epsilon(y)$ vanish and we are left with

$$(3.4.7) \quad |f(x) - Q_\epsilon(x) * f(x)| \leq \text{Const.} \cdot \text{Max}_{|y-x| \leq \epsilon} |f^{(r+1)}(y)| \cdot \epsilon^{r+1},$$

i.e., $f_\epsilon(x)$ converges to $f(x)$ with order $r+1$ as $\epsilon \rightarrow 0$. Moreover, $f_\epsilon(x)$ is as smooth as $\phi(x)$ is, since

$$(3.4.8) \quad f_\epsilon(x) = \int_y \frac{1}{\epsilon} Q\left(\frac{x-y}{\epsilon}\right) f(y) dy$$

has many bounded derivatives as Q has, i.e., starting with differentiable function f of order $r+1$ in the neighborhood of x , we end up with regularized function $f_\epsilon(x)$ in C^s , $s > r$.

Example: For C^∞ regularization – choose a unit mass C^∞ kernel, see Figure 6,

$$(3.4.9a) \quad \phi(x) = \begin{cases} Q_0 e^{-\frac{1}{1-x^2}}, & |x| < 1 \\ 0, & |x| \geq 1 \end{cases} \quad \text{with } Q_0 \text{ such that } \int \phi(x) dx = 1.$$

Then $f_\epsilon(x) = Q_\epsilon(x) * f(x)$ is a C^∞ regularization of $f(x)$ with first order convergence rate

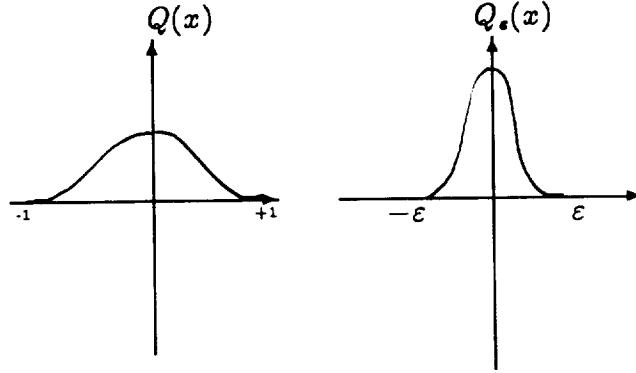


Figure 6:

$$(3.4.9b) \quad |f(x) - f_\epsilon(x)| \leq \text{Const.} \cdot \text{Max}_{|y-x| \leq \epsilon} |f'(y)| \cdot \epsilon \rightarrow 0.$$

To increase the order of convergence, we require more vanishing moments which yield more oscillatory kernels as in Figure 7.

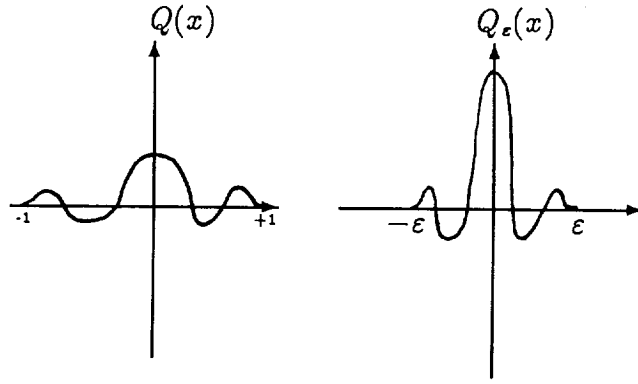


Figure 7:

We note that this smoothing process is purely local – it involves ε -neighboring values of C^{r+1} function f , in order to yield a C^r -regularized function $f_\varepsilon(x)$ with $f_\varepsilon(x) \rightarrow f(x)$. The convergence rate here is $r + 1$.

We can also achieve local regularization with spectral convergence. To this end set

$$(3.4.10a) \quad Q_N(x) = \rho(x)D_N(x)$$

where $\rho(x)$ is a C^∞ -function supported on $(-\varepsilon, \varepsilon)$ such that, see Figure 8,

$$(3.4.10b) \quad \rho(0) = 1.$$

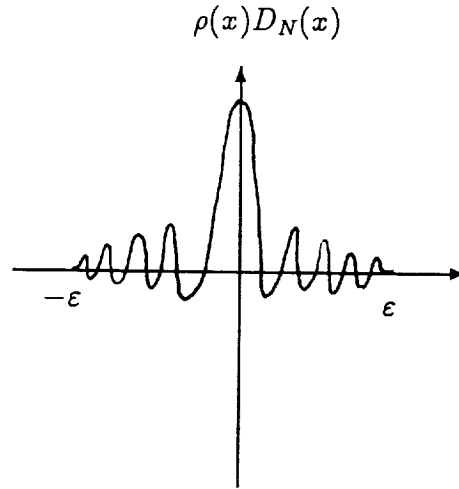


Figure 8:

Consider now:

$$(3.4.11) \quad f_N(x) = Q_N * f(x).$$

Then $f_N(x)$ is C^∞ because Q_N is; and the convergence rate is spectral, since by (1.1.25)

$$(3.4.12a) \quad \begin{aligned} f(x) - Q_N * f(x) &= f(x) - \int_{|y| \leq \epsilon} D_N(y) \rho(y) f(x-y) dy \\ &= f(x) - \rho(y) f(x-y)|_{y=0} + \text{residual} \end{aligned}$$

where

$$(3.4.12b) \quad |\text{residual}| \leq \text{Const.} \|\rho(y) f(x-y)\|_{H^s(-\epsilon, \epsilon)} \frac{1}{N^{s-1}} \text{ for any } s > 0.$$

Thus, the convergence rate is as fast as the local smoothness of f permits; of course with $\rho \equiv 1$ we obtain the C^∞ -regularization due to the spectral projection. The role of ρ was to localize this process of spectral smoothing.

We can as easily implement such smoothing in the Fourier space: For example, with the heat kernel we have

$$(3.4.13) \quad \hat{u}(k, t) = e^{-ak^2 t} \hat{f}(k)$$

so that $\hat{u}(k, t)$ for any $t > 0$ decay faster than exponential and hence $u(x, t > 0)$ belong to H^s for any s (and by Sobolev imbeddings, therefore, is in C^∞ and in fact analytic). In general we apply,

$$(3.4.14) \quad f_\epsilon(x) = \sum_{k=-\infty}^{\infty} \hat{Q}_\epsilon(k) \hat{f}(k) e^{ikx}$$

such that for $f_\epsilon(x)$ to be in H^s we require

$$(3.4.15a) \quad \sum_{k=-\alpha}^{\infty} (1 + |k|^2)^s |\hat{Q}_\epsilon(k)|^2 |\hat{f}(k)|^2 \leq \text{Const.}$$

and $r+1$ order of convergence follows with

$$(3.4.15b) \quad |\hat{\phi}_\epsilon(k) - 1| \leq \text{Const.} (\epsilon k)^{r+1}.$$

Indeed, (3.4.15b) implies

$$(3.4.16) \quad \begin{aligned} |f(x) - f_\epsilon(x)| &\leq \text{Const.} \epsilon^{r+1} \sum_{k=-\alpha}^{\infty} |k|^{r+1} |\hat{f}(k)| e^{ikx} \\ &\leq \text{Const.} \text{Max}|f^{(r+1)}| \cdot \epsilon^{r+1}. \end{aligned}$$

Note: Since $\hat{\phi}_\epsilon(k) \downarrow 0$ we can deal with any unbounded f by splitting $\sum_{|k| \leq |k_0|} + \sum_{|k| > |k_0|}$. To obtain spectral accuracy we may use

$$(3.4.17a) \quad \hat{Q}_N(k) = \begin{cases} \equiv 1, & |k| < \frac{N}{2} \\ \sim \text{smoothly decay to zero} & \frac{N}{2} \leq |k| \leq N \end{cases}$$

Clearly $Q_N * f(x)$ is C^∞ and as in Section 1

$$(3.4.17b) \quad |f(x) - Q_N * f(x)| \leq \sum_{|k| > N_\epsilon} |\hat{f}(k) e^{ikx}| \leq \text{Const.} \|f\|_{H^s} \cdot \frac{1}{N^{s-1}}.$$

We emphasize that this kind of smoothing in the Fourier space need not be local; rather $Q_\epsilon(x)$ or $\phi_N(x)$ are negligibly small away from a small interval centered around the origin depending on ϵ or $\frac{1}{N}$. (This is due to the uncertainty principle.)

The smoothed version of the pseudospectral approximation of (3.2.16) reads

$$(3.4.18) \quad \frac{\partial v_N}{\partial t} = \psi_N(a(x, t)) \frac{\partial}{\partial x} (Q * v_N)$$

i.e., in each step we smooth the solution either in the real space (convolution) or in the Fourier space (cutting high modes).⁵ We claim that this smoothed version is stable hence convergent under very mild assumptions on the smoothing kernel $Q_N(x)$. Specifically, (3.4.18) amounts in the Fourier space, compare (3.2.3)

$$(3.4.19) \quad \frac{\partial \tilde{v}}{\partial t} = \tilde{A}_N \Lambda Q_N \tilde{v}.$$

The real part of the matrix in question is given by

$$(3.4.20a) \quad [\text{Re} \tilde{A}_N \Lambda Q_N]_{kj} = i(\lambda_k - \lambda_j) \sum_p \hat{a}[k - j + p(2N + 1)], \quad -N \leq k, j \leq N$$

where $\Lambda Q_N = \text{diag}_k(i\lambda_k)$

$$(3.4.20b) \quad i\lambda_k = ik\hat{Q}_N(k)$$

is interpreted as the smoothed differentiation operator. Now, looking at (3.4.20a) we note:

1. For $p = 0$ we are back at the spectral analysis, compare (3.1.12), (3.1.13) and the real part of the matrix in (3.4.20a) – the aliasing free one – is bounded.
2. We are left with $|p| = 1$: in the unsmoothed version, these terms were unbounded since $|\lambda_k - \lambda_j| \uparrow \infty$ as $k \downarrow -N$ or $j \uparrow N$. With the smoothed version, these terms are bounded (and stability follows), provided we have

⁵Either one can be carried out efficiently by the FFT.

$$(3.4.21) \quad |\lambda_k = ik\hat{Q}_N(k)| \xrightarrow{|k| \uparrow N} 0.$$

For example, consider the smoothing kernel $Q_N(x)$ where

$$(3.4.22a) \quad \hat{Q}_N(k) = \frac{\sin kh}{kh}, \quad h = \frac{2\pi}{2N+1}.$$

This yields the smoothed differentiation symbols

$$(3.4.22b) \quad \lambda_k = i \sin \frac{kh}{h}$$

which corresponds to the second order center differencing in (1.2.36); stability is immediate by (3.4.21) for

$$|\lambda_k| \equiv \left| \frac{\sin \frac{2\pi k}{2N+1}}{\frac{2\pi k}{2N+1}} \right| \xrightarrow{|k| \uparrow N} 0.$$

Yet, this kind of smoothing reduces the overall spectral accuracy to a second one; a fourth order smoothing will be

$$(3.4.23a) \quad \lambda_k = i \frac{1}{3} \left[4 \sin \frac{kh}{h} - \sin \frac{2kh}{2h} \right], \quad \hat{Q}_N(k) = \frac{\lambda_k}{ik}$$

or

$$(3.4.23b) \quad \lambda_k = \frac{6i}{h} \frac{\sin kh}{4 + 2 \cos kh}, \quad \hat{Q}_N(k) = \frac{\lambda_k}{ik}.$$

In general, the accuracy is determined by the low modes while stability has to do with high ones. To entertain spectral accuracy we may consider smoothing kernels other than trigonometric polynomials (\equiv finite difference), but rather, compare (3.4.17)

$$(3.4.24) \quad \hat{Q}_N(k) = \begin{cases} \equiv 1, & |k| \leq \frac{N}{2} \\ \sim \text{smoothly decay to zero} & \frac{N}{2} < |k| \leq N. \end{cases}$$

An increasing portion of the spectrum is differentiated exactly which yields spectral accuracy; the highest modes are not amplified because of the smoothing effect in this part of the spectrum.

We close this section noting that if some dissipation is present in the differential model to begin with, e.g., with the parabolic equation

$$(3.4.25) \quad \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(a(x, t) \frac{\partial u}{\partial x} \right), \quad a(x, t) \geq \alpha > 0,$$

then stability follows with no extra smoothing. The parabolic dissipation compensates for the loss of "one derivative" if first order terms are present. To see this we proceed as follows: multiply

$$(3.4.26) \quad \frac{\partial v_N}{\partial t}(x_\nu, t) = \frac{\partial}{\partial x} \left[a(x_\nu, t) \frac{\partial v_N}{\partial x}(x_\nu, t) \right]$$

by $v_N(x_\nu, t)$ and sum to obtain

$$(3.4.27) \quad \frac{1}{2} \frac{d}{dt} \sum_{\nu} v_N^2(x_\nu, t) = \sum_{\nu} v_N(x_\nu, t) \frac{\partial}{\partial x} (a(x_\nu, t) \frac{\partial v_N}{\partial x}(x_\nu, t)).$$

Suppressing excessive indices, $v_N(x_\nu, t) \equiv v_\nu(t)$, we have for the RHS of (3.4.27)

$$(3.4.28) \quad \sum_{\nu} v_{\nu} \frac{\partial}{\partial x} \left(a_{\nu}(t) \frac{\partial v_{\nu}}{\partial x} \right) = \frac{1}{2} \sum_{\nu} \frac{\partial}{\partial x} \left(a_{\nu}(t) \frac{\partial v_{\nu}^2}{\partial x} \right) - \sum_{\nu} a_{\nu}(t) \left(\frac{\partial v_{\nu}}{\partial x} \right)^2.$$

Now, the first sum on the right gives us the usual loss of one derivative and the second are compensates with gain of such quantity. Petrowski type stability (gain of derivatives) follows. We shall only sketch the details here. Starting with the first term on the right of (3.4.28) we have

$$(3.4.29) \quad \frac{1}{2} \sum \frac{\partial}{\partial x} \left(a_{\nu} v_{\nu} \frac{\partial v_{\nu}}{\partial x} \right) = \frac{1}{2} \int \frac{\partial}{\partial x} [\cdots] + \frac{1}{2} \cdot [\text{aliasing errors}]$$

while for the second term

$$(3.4.30) \quad - \sum a_{\nu}(t) \left(\frac{\partial v_{\nu}}{\partial x} \right)^2 \leq -\alpha \int \left[\frac{\partial v_N}{\partial x}(x, t) \right]^2 dx$$

and this last term dominates the RHS of (3.4.29).

APPENDIX

A.1. Fourier Collocation with Even Number of Gridpoints

We assume $w(x)$ is known at

$$(A.1.1) \quad w_\nu = w(x_\nu) \quad x_\nu = r + \nu h \quad \nu = 0, 1, \dots, 2N$$

with $h \equiv \frac{2\pi}{2N} = \frac{\pi}{N}$ and $0 \leq r < h$. We use the trapezoidal nodes

$$(A.1.2) \quad \tilde{w}(k) = \frac{1}{2\pi} \sum_{\nu=0}^{2N} w_\nu e^{-ikx_\nu} h = \frac{1}{2N} \sum_{\nu=0}^{2N-1} w_\nu e^{-ikx_\nu}$$

to obtain the pseudospectral approximation

$$(A.1.3) \quad \psi_N w = \sum_{k=-N}^N \tilde{w}(k) e^{ikx}.$$

Note: We now have only $2N$ pieces of discrete data at the different $2N$ grid points $x_0, x_1, \dots, x_{2N-1}$ and they correspond to $2N$ waves, as we have a “silent” last mode, i.e., with $r = 0$, $k = N$, $Im[e^{ikx}]_{x=x_\nu} = i \sin \nu\pi = 0$. This is a projection, since in view of (A.1.3) $\psi_N w$ is the interpolant of $w(x)$ at $x = x_\nu$:

$$(A.1.4) \quad \begin{aligned} \psi_N w(x)|_{x=x_\mu} &= \sum_{k=-N}^N \left[\frac{1}{2N} \sum_{\nu=0}^{2N-1} w(x_\nu) e^{-ikx_\nu} \right] e^{ikx_\mu} \\ &= \sum_{\nu=0}^{2N-1} w(x_\nu) \cdot \frac{1}{2N} \sum_{k=-N}^N e^{ik(\mu-\nu)h} = w(x_\mu). \end{aligned}$$

The aliasing relation in this case reads – compare (1.2.7)

$$(A.1.5) \quad \tilde{w}(k) = \sum_{p=-\infty}^{\infty} e^{ip2N\tau} \hat{w}(k + 2pN)$$

and spectral convergence follow – compare with (1.2.16)

$$(A.1.6) \quad \|A_N w(x)\|_{H^s} \leq \text{Const}_s \cdot \|T_N w(x)\|_{H^s}, \quad s > \frac{1}{2}.$$

In the usual sin-cos formulation it takes the form

$$(A.1.7) \quad \psi_N w = \sum_{k=0}^N \tilde{a}_k \cos kx + \tilde{b}_k \sin kx, \quad \begin{bmatrix} \tilde{a}_k \\ \tilde{b}_k \end{bmatrix} = \frac{1}{N} \sum_{\nu=0}^{2N-1} w(x_\nu) \begin{bmatrix} \cos kx_\nu \\ \sin kx_\nu \end{bmatrix}, \quad 0 \leq k \leq N.$$

Noting that $\tilde{b}_N = 0$ we have $2N$ free parameters $\tilde{a}_0, \{\tilde{a}_k, \tilde{b}_k\}_{k=1}^{N-1}$ and \tilde{a}_N to match our data at $\{x_\nu\}_{\nu=0}^{2N-1}$.



Report Documentation Page

1. Report No. NASA CR-187443 ICASE Interim Report 14		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle SPECTRAL METHODS FOR TIME DEPENDENT PROBLEMS				5. Report Date September 1990	
				6. Performing Organization Code	
7. Author(s) Eitan Tadmor				8. Performing Organization Report No. Interim Report No. 14	
				10. Work Unit No. 505-90-21-01	
9. Performing Organization Name and Address Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23665-5225				11. Contract or Grant No. NAS1-18605	
				13. Type of Report and Period Covered Contractor Report	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Langley Research Center Hampton, VA 23665-5225				14. Sponsoring Agency Code	
15. Supplementary Notes Langley Technical Monitor: Richard W. Barnwell Interim Report					
16. Abstract This short review on spectral approximations for time-dependent problems consists of three parts. In part I we discuss some basic ingredients form the spectral Fourier and Chebyshev approximation theory. Part II contains a brief survey on hyperbolic and parabolic time-dependent problems which are dealt with both the energymethod and the related Fourier analysis. In part III we combine the ideas presented in the first two parts, in our study of accuracy stability and convergence of the spectral Fourier approximation to time-dependent problems.					
17. Key Words (Suggested by Author(s)) Fourier approximation, Chebyshev approximation, spectral accuracy, hyperbolic and parabolic systems, stability, convergence			18. Distribution Statement 64 - Numerical Analysis Unclassified - Unlimited		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of pages 68	
				22. Price A04	

